

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 06:13:09 ; Search time 51 Seconds
(without alignments)
487.075 Million cell updates/sec

Title: US-09-724-000A-5

Perfect score: 442
Sequence: 1 MRLVLVSLLCILLICRSIF.....PCKLEPRRLMVVPGALPQV 81

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09724000/runat_16052003_145501_17219/app_query.fasta_1.263
-DB=Issued Patents NA -QFMT=fastrap -SUFFIX=rni -MINMATCH=0.1 -LOFPCI=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09724000 @CGN 1 1 40 @runat_16052003_145501_17219 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WAEI TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	85	19.2	1688	2	US-08-649-6198-2
C 2	85	19.2	1979	2	US-08-649-6198-1
C 3	85	19.2	2030	1	US-08-330-518-1
C 4	85	19.2	2030	1	US-08-330-283-1
C 5	85	19.2	2030	2	US-08-646-248-1
C 6	85	19.2	2030	5	PCT-US95-13924-1
C 7	85	19.2	2030	5	PCT-US95-13931-1
C 8	78	17.6	1813	5	PCT-US94-12883-3
C 9	78	17.6	1898	1	US-08-342-411A-1
C 10	74.5	16.9	48974	4	US-08-920-422-17
C 11	73.5	16.6	1600	4	US-09-029-027B-1
C 12	73.5	16.6	2542	4	US-09-305-681-5

13	72	16.3	1493	2	US-08-752-307B-6	Sequence 6, Appli
14	72	16.3	1493	4	US-09-707-802-6	Sequence 6, Appli
15	72	16.3	1493	4	US-09-991-326-6	Sequence 6, Appli
16	72	16.3	2244	4	US-09-354-151-1	Sequence 1, Appli
17	72	16.3	4086	1	US-08-313-181-1	Sequence 1, Appli
C 18	71.5	16.2	888	4	US-09-215-694-30	Sequence 30, Appli
19	71.5	16.2	1736	4	US-09-162-524-2	Sequence 2, Appli
C 20	71.5	16.2	1798	4	US-09-305-681-1	Sequence 1, Appli
C 21	71.5	16.2	2416	1	US-08-592-126-61	Sequence 61, Appli
22	71.5	16.2	31328	4	US-09-215-694-19	Sequence 19, Appli
23	71	16.1	3807	2	US-08-816-755-1	Sequence 1, Appli
24	71	16.1	3807	4	US-09-090-673-1	Sequence 1, Appli
25	70.5	16.0	2219	1	US-08-606-322-1	Sequence 1, Appli
26	70.5	16.0	3502	2	US-08-724-394A-16	Sequence 16, Appli
27	70.5	16.0	4040	4	US-08-664-962B-1	Sequence 1, Appli
28	70.5	16.0	4040	4	US-09-311-743-1	Sequence 1, Appli
29	70	15.8	2743	1	US-08-396-479B-3	Sequence 3, Appli
30	70	15.8	2743	1	US-08-818-823-3	Sequence 3, Appli
31	70	15.8	2749	2	US-08-144-981A-1	Sequence 1, Appli
32	70	15.8	2751	3	US-09-037-190-45	Sequence 45, Appli
33	70	15.8	2751	3	US-09-037-192-45	Sequence 45, Appli
34	70	15.8	2751	3	US-09-037-143-45	Sequence 45, Appli
35	70	15.8	2751	4	US-09-049-691-45	Sequence 45, Appli
36	70	15.8	2751	4	US-08-260-174-45	Sequence 45, Appli
37	70	15.8	2751	4	US-09-338-128A-45	Sequence 45, Appli
38	70	15.8	2853	5	PCT-US94-07297-36	Sequence 36, Appli
39	70	15.8	2853	5	US-09-232-346-45	Sequence 45, Appli
40	69.5	15.7	1110	4	US-08-462-509B-3	Sequence 3, Appli
41	69.5	15.7	1110	5	PCT-US95-05616-3	Sequence 3, Appli
42	69.5	15.7	1654	3	US-08-991-426-3	Sequence 3, Appli
43	69.5	15.7	1654	3	US-09-143-470-3	Sequence 3, Appli
44	69.5	15.7	2959	1	US-08-408-188A-1	Sequence 1, Appli
C 45	69.5	15.7	2959	1	US-08-408-188A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-08-649-619B-2/C

; Sequence 2, Application US/08649619B

; Patent No. 5871916

; GENERAL INFORMATION:

; APPLICANT: NAKAMURA, YUSUKE

; TITLE OF INVENTION: ECDN PROTEIN AND DNA

; TITLE OF INVENTION: ENCODING THE SAME

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLYNN, THIEL, BOUTELL &

; ADDRESSEE: TANIS, P.C.

; STREET: 2026 Rambling Road

; CITY: Kalamazoo

; STATE: Michigan

; COUNTRY: USA

; ZIP: 49008-1699

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb

; MEDIUM TYPE: storage

; COMPUTER: IBM PC/XT/AT Compatible

; OPERATING SYSTEM: MS-DOS 5.0

; SOFTWARE: WordPerfect 5.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/649,619B

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP6-226270

; FILING DATE: 21-SEPT-1994

; APPLICATION NUMBER: PCT/JP95/01909

; FILING DATE: 21-SEPT-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Terryence F. Chapman

; REGISTRATION NUMBER: 32549

REFERENCE/DOCKET NUMBER: Furuya Case 1343
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1688
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: Homosapiens
US-08-649-619B-2

Alignment Scores:
Pred. No.: 1.65 Length: 1688
Score: 85.00 Matches: 26
Percent Similarity: 47.76% Conservative: 6
Best Local Similarity: 38.81% Mismatches: 17
Query Match: 19.23% Indels: 18
DB: 2 Gaps: 4

US-09-724-000A-5 (1-81) x US-08-649-619B-2 (1-1688)

Qy 22 ThrGluGlyLysArgProAlaLysAla-----TrrSerGlyArgThrArgLeu 39
Db 692 ACTGGGGTCTCGCCGCGCCAGGCGGTGACTTTGGGCTGTCGAGAGGCGTT 633

Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisValArg 58
Db 632 TGTTCACGTGCGAGTTGGCGCGCCACCACTGCTGGATCATTA----- 591

Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 590 -----GTTCTTGAGCGCTGTTAGTGGACACCCCTCGCTTCCCGGAGC 546

Qy 75 ProGlyAlaLeuProGlnVal 81
Db 545 CCTGGCTGCCTGCCTCAGATC 525

RESULT 2

US-08-649-619B-1/c
Sequence 1, Application US/08649619B
Patent No. 5671916

GENERAL INFORMATION:
APPLICANT: NAKAMURA, YUSUKE
APPLICANT: SAITO, HIROKO
TITLE OF INVENTION: ECDN PROTEIN AND DNA
TITLE OF INVENTION: ENCODING THE SAME
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLYNN, THIEL, BOUTELL &
STREET: 2026 Rambling Road
CITY: Kalamazoo
STATE: Michigan
COUNTRY: USA
ZIP: 49008-1699

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inches, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM PC/XT/AT Compatible
OPERATING SYSTEM: MS-DOS 5.0
SOFTWARE: Wordperfect 5.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/649,619B
FILING DATE:
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP6-226270
FILING DATE: 21-SEPT-1994
APPLICATION NUMBER: PCT/JP95/01909

FILING DATE: 21-SEPT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Terrylene F. Chapman
REGISTRATION NUMBER: 32549
REFERENCE/DOCKET NUMBER: Furuya Case 1343
TELECOMMUNICATION INFORMATION:
TELEPHONE: (616) 381-1156
TELEFAX: (616) 381-5465
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1979
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE: Homosapiens
IMMEDIATE SOURCE:
LIBRARY: Human mammary gland cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 206..1591
IDENTIFICATION METHOD: experimental examination
US-08-649-619B-1

Alignment Scores:
Pred. No.: 2.06 Length: 1979
Score: 85.00 Matches: 26
Percent Similarity: 47.76% Conservative: 6
Best Local Similarity: 38.81% Mismatches: 17
Query Match: 19.23% Indels: 18
DB: 2 Gaps: 4

US-09-724-000A-5 (1-81) x US-08-649-619B-1 (1-1979)

Qy 22 ThrGluGlyLysArgProAlaLysAla-----TrrSerGlyArgThrArgLeu 39
Db 983 ACTGGGGTCTCGCCGCGCCAGGCGGTGACTTTGGGCTGTCGAGAGGCGTT 924

Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisValArg 58
Db 923 TGTTCACGTGCGAGTTGGCGCGCCACCACTGCTGGATCATTA----- 882

Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 881 -----GTTCTTGAGCGCTGTTAGTGGACACCCCTCGCTTCCCGGAGC 837

Qy 75 ProGlyAlaLeuProGlnVal 81
Db 836 CCTGGCTGCCTGCCTCAGATC 816

RESULT 3

US-08-330-518-1/c
Sequence 1, Application US/08330518
Patent No. 5607967

GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: USE OF RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,518
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DOCKET NUMBER: 19316
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-330-518-1

Alignment Scores:
Pred. No.: 2.13 Length: 2030
Score: 85.00 Matches: 26
Percent Similarity: 47.76% Conservative: 6
Best Local Similarity: 38.81% Mismatches: 17
Query Match: 19.23% Indels: 18
DB: 1 Gaps: 4

US-09-724-000A-5 (1-81) x US-08-330-518-1 (1-2030)

Qy 22 ThrGlugLyysArgArgProAlaLysAla-----TTPSerGlyArgArgThrArgLeu 39
Db 1022 ACTGGGGGTCTGCCGCCAGGGCCAGGCGTGACTTTGGGCTGTCGAGAGAGCGTT 963
Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAnLeuLysGlyHisHisValArg 58
Db 962 TGTTCACACTGAGTTGGCGCCACCACTGCTGATCATTA----- 921
Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 920 -----GTTCTTGAGCGCTGTTAGCTGGACACCCCTCGCCTTCCCGGAGC 876
Qy 75 ProGlyAlaLeuProGlnVal 81
Db 875 CTGGCTGCCTGCCTCAGATC 855

RESULT 4
US-08-330-283-1/c
Sequence 1, Application US/08330283
Patent No. 5679518
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,283
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dolan, Catherine A.
REGISTRATION NUMBER: 36,502
REFERENCE/DOCKET NUMBER: 19327
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-4283
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-330-283-1

Alignment Scores:
Pred. No.: 2.13 Length: 2030
Score: 85.00 Matches: 26
Percent Similarity: 47.76% Conservative: 6
Best Local Similarity: 38.81% Mismatches: 17
Query Match: 19.23% Indels: 18
DB: 1 Gaps: 4

US-09-724-000A-5 (1-81) x US-08-330-283-1 (1-2030)

Qy 22 ThrGlugLyysArgArgProAlaLysAla-----TTPSerGlyArgArgThrArgLeu 39
Db 1022 ACTGGGGGTCTGCCGCCAGGGCCAGGCGTGACTTTGGGCTGTCGAGAGAGCGTT 963
Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAnLeuLysGlyHisHisValArg 58
Db 962 TGTTCACACTGAGTTGGCGCCACCACTGCTGATCATTA----- 921
Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 920 -----GTTCTTGAGCGCTGTTAGCTGGACACCCCTCGCCTTCCCGGAGC 876
Qy 75 ProGlyAlaLeuProGlnVal 81
Db 875 CTGGCTGCCTGCCTCAGATC 855

RESULT 5
US-08-646-248-1/c
Sequence 1, Application US/08646248
Patent No. 5939322
GENERAL INFORMATION:
APPLICANT: Friedman, Eitan
APPLICANT: Holloway, M. Katharine
APPLICANT: Rodan, Gideon
APPLICANT: Rutledge, Su Jane
APPLICANT: Schmidt, Azriel
APPLICANT: Vogel, Robert
TITLE OF INVENTION: METHOD FOR FINDING RECEPTOR POTENTIATORS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Merck & Co., Inc.
STREET: 126 East Lincoln Avenue
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13931
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Quagliato, Carol S.
REGISTRATION NUMBER: 35,330
REFERENCE/DOCKET NUMBER: 19316 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-3809
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2030 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US95-13931-1

Alignment Scores:
Pred. No.: 2,13 Length: 2030
Score: 85.00 Matches: 26
Percent Similarity: 47.76% Conservatives: 6
Best Local Similarity: 38.81% Mismatches: 17
Query Match: 19.23% Indels: 18
DB: 5 Gaps: 4

US-09-724-000A-5 (1-81) x PCT-US95-13931-1 (1-2030)

Qy 22 ThrGlucLyAsArgProAlaLysAla-----TrpSerGlyArgArgThrArgLeu 39
Db 1022 ACTGGGGTCTGCGCCCGAGGCGCGTGGACTTTGGGCTGTCGAGAGAGCGTT 963
Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
Db 962 TGTTCAGTCTGAGTGGCGCCCGACCACTGCTGCATCATTA----- 921
Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 920 -----GTTCTTGAGCGCGTGTAGTGTGACACCCCTGCTCCCGGAGC 876
Qy 75 ProGlyAlaLeuProGlnVal 81
Db 875 CTGGCTGCTGCTCAGATC 855

RESULT 8
PCT-US94-12883-3/c
Sequence 3, Application PC/TUS9412883
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: UBIQUITOUS NUCLEAR RECEPTOR: COMPOSITIONS AND
METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States of America
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: PatentIn Release #1.0, Version
#1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/12883
FILING DATE: Concurrently Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/152,003
FILING DATE: 10-NOV-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: BARBARA S. KITCHELL
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: ARCD154P--
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
TELEX: 79-0924
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1813 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
PCT-US94-12883-3

Alignment Scores:
Pred. No.: 11,7 Length: 1813
Score: 78.00 Matches: 25
Percent Similarity: 46.27% Conservatives: 6
Best Local Similarity: 37.31% Mismatches: 18
Query Match: 17.65% Indels: 18
DB: 5 Gaps: 4

US-09-724-000A-5 (1-81) x PCT-US94-12883-3 (1-1813)

Qy 22 ThrGlucLyAsArgProAlaLysAla-----TrpSerGlyArgArgThrArgLeu 39
Db 760 ACTGGGGTCTGCGCCCGAGGCGCGTGGACTTTGGGCTGTCGAGAGAGCGTT 701
Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisHisValArg 58
Db 700 TGTTCAGTCTGAGTGGCGCCCGACCACTGCTGCATCATTA----- 659
Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 658 -----GTTCTTGAGCGCGTGTAGTGTGACACCCCTGCTCCCGGAGC 614
Qy 75 ProGlyAlaLeuProGlnVal 81
Db 613 CTGGCTGCTGCTCAGATC 593

RESULT 9
US-08-342-411A-1/c
Sequence 1, Application US/08342411A
Patent No. 5639616
GENERAL INFORMATION:
APPLICANT: LIAO, Shutsung
TITLE OF INVENTION: UBIQUITOUS NUCLEAR RECEPTOR:
COMPOSITIONS AND METHODS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/305.681
FILING DATE:
CLASSIFICATION:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2542 base pairs
TYPE: nucleotide
STRANDEDNESS: single strand
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-09-305-681-5

Alignment Scores:
Pred. No.: 61.6 Length: 2542
Score: 73.50 Matches: 22
Percent Similarity: 44.64% Conservative: 3
Best Local Similarity: 39.29% Mismatches: 22
Query Match: 16.63% Indels: 9
DB: 4 Gaps: 4

US-09-724-000A-5 (1-81) x US-09-305-681-5 (1-2542)

Qy 24 GlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCys----- 41
Db 1400 GGGGACGCGCGCTGCTGCGGTACGAGGGGT-----CCTCGCTTATGCTGTCAATAA 1347
Qy 42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArg-----Leu 59
Db 1346 CATCGGGCTTAAGCCCCCACCAGCACGTAGCCAGGGGCCATTGCCATAGAGGATGGAGG 1287
Qy 60 CysLysProCysLysLeuGluPro-----GluProArgLeuTrp 72
Db 1286 TGTAGGACTTGTGTCTGAGGCTTGTCTGGGGCCCTGCTGGGGCCGACAGCCGAATGG 1239

RESULT 13

US-08-752-307B-6
Sequence 6, Application US/08752307B
Patent No. 5952171
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearing, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/752.307B
FILING DATE: 19-NOV-1996
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 09404/020001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906

TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 99...1493
US-08-752-307B-6

Alignment Scores:
Pred. No.: 43.8 Length: 1493
Score: 72.00 Matches: 18
Percent Similarity: 39.39% Conservative: 8
Best Local Similarity: 27.27% Mismatches: 24
Query Match: 16.29% Indels: 16
DB: 2 Gaps: 3

US-09-724-000A-5 (1-81) x US-08-752-307B-6 (1-1493)
Qy 27 ArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCysCys----- 41
Db 1174 AGGCCATCTCCATCC--GTGGGCTCAGCAACGAGACGCTCTCATCACCTCGGCCGAGA 1230
Qy 42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHis----- 55
Db 1231 AGAGCCATTCGGGGCTTACCAGTGTCTACCGAAGCCGACCCAGGACT 1290
Qy 56 -----HisValArgLeuCysLysProCysLysLeuGluProArgLeu 71
Db 1291 TTGCCATCATTTGCACTTGAGGATGCGACGCCCGCCGATCGTCTGCTCTTACGCGAGAAGG 1350
Qy 72 TrpValValProGlyAla 77
Db 1351 TGGTCAACCCCGGGGAGC 1368

RESULT 14

US-09-707-802-6
Sequence 6, Application US/09707802
Patent No. 6391586
GENERAL INFORMATION:
APPLICANT: McCarthy, Sean A.
APPLICANT: Gearing, David P.
APPLICANT: Levinson, Douglas A.
TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/707.802
FILING DATE: 07-NO. 6391586-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 09404/020001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 99...1493

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-707-802-6

Alignment Scores:
Pred. No.: 43.8 Length: 1493
Score: 72.00 Matches: 18
Percent Similarity: 39.39% Conservative: 8
Best Local Similarity: 27.27% Mismatches: 24
Query Match: 16.29% Indels: 16
DB: 4 Gaps: 3

US-09-724-000A-5 (1-81) x US-09-707-802-6 (1-1493)

Qy 27 ArgProAlaLysAlaTrpSerGlyArgThrArgLeuCysCys----- 41
Db 1174 AGGCCATCTCCATCC---GTGGGCTCAGCAACGAGACGCTGCTCATCACCTCGGCCCCAGA 1230
Qy 42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHis----- 55
Db 1231 AGAGCCATTCCGGGCGCTACCAAGTCTGCTACCCGCAAGCCAGACCCGCCAGGACT 1290
Qy 56 -----HisValArgLeuCysLysProCysLysLeuGluProGluProArgLeu 71
Db 1291 TTGCCATCATTTGACCTTGAGGATGCACGCCCGCATCGTCTCGTCTTCAGCGAGAAGG 1350
Qy 72 TrpValValProGlyAla 77
Db 1351 TGGTCAACCCCGGGGAGC 1368

RESULT 15

US-09-991-326-6

Sequence 6, Application US/09991326

Patent No. 6395872

GENERAL INFORMATION:

APPLICANT: McCarthy, Sean A.

Gearing, David P.

Levinson, Douglas A.

TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES

ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson, P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/991,326

FILING DATE: 21-Nov-1996

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/752,307

FILING DATE: 19-Nov-1996

ATTORNEY/AGENT INFORMATION:

NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 09404/020002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
TELEX: 200154

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 1493 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 99...1493

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-991-326-6

Alignment Scores:
Pred. No.: 43.8 Length: 1493
Score: 72.00 Matches: 18
Percent Similarity: 39.39% Conservative: 8
Best Local Similarity: 27.27% Mismatches: 24
Query Match: 16.29% Indels: 16
DB: 4 Gaps: 3

US-09-724-000A-5 (1-81) x US-09-991-326-6 (1-1493)

Qy 27 ArgProAlaLysAlaTrpSerGlyArgThrArgLeuCysCys----- 41
Db 1174 AGGCCATCTCCATCC---GTGGGCTCAGCAACGAGACGCTGCTCATCACCTCGGCCCCAGA 1230
Qy 42 HisArgValProSerProAsnSerThrAsnLeuLysGlyHis----- 55
Db 1231 AGAGCCATTCCGGGCGCTACCAAGTCTGCTACCCGCAAGCCAGACCCGCCAGGACT 1290
Qy 56 -----HisValArgLeuCysLysProCysLysLeuGluProGluProArgLeu 71
Db 1291 TTGCCATCATTTGACCTTGAGGATGCACGCCCGCATCGTCTCGTCTTCAGCGAGAAGG 1350
Qy 72 TrpValValProGlyAla 77
Db 1351 TGGTCAACCCCGGGGAGC 1368

Search completed: May 17, 2003, 13:36:54

Job time : 65 secs

Result No.	Score	Query Match	Length	DB ID	Description
C 1	442	100.0	342	9	Sequence 3006, App
C 2	442	100.0	379	9	Sequence 671, App
C 3	442	100.0	379	9	Sequence 671, App
C 4	442	100.0	379	9	Sequence 671, App
C 5	442	100.0	379	9	Sequence 671, App
C 6	442	100.0	379	9	Sequence 671, App
C 7	442	100.0	379	9	Sequence 671, App
C 8	442	100.0	379	9	Sequence 671, App
C 9	442	100.0	379	9	Sequence 671, App
C 10	442	100.0	379	9	Sequence 671, App
C 11	442	100.0	379	9	Sequence 671, App
C 12	442	100.0	379	9	Sequence 671, App
C 13	442	100.0	379	9	Sequence 671, App
C 14	442	100.0	379	9	Sequence 671, App
C 15	442	100.0	379	9	Sequence 671, App
C 16	442	100.0	379	9	Sequence 671, App
C 17	442	100.0	379	9	Sequence 671, App
C 18	442	100.0	379	9	Sequence 671, App
C 19	442	100.0	379	9	Sequence 671, App
C 20	442	100.0	379	9	Sequence 671, App
C 21	442	100.0	379	9	Sequence 671, App
C 22	442	100.0	379	9	Sequence 671, App
C 23	442	100.0	379	9	Sequence 671, App
C 24	442	100.0	379	9	Sequence 671, App
C 25	442	100.0	379	9	Sequence 671, App
C 26	442	100.0	379	9	Sequence 671, App
C 27	442	100.0	379	9	Sequence 671, App
C 28	442	100.0	379	9	Sequence 671, App
C 29	442	100.0	379	9	Sequence 671, App
C 30	442	100.0	379	9	Sequence 671, App
C 31	442	100.0	379	9	Sequence 671, App
C 32	442	100.0	379	9	Sequence 671, App
C 33	442	100.0	379	9	Sequence 671, App
C 34	442	100.0	379	9	Sequence 671, App
C 35	442	100.0	379	9	Sequence 671, App
C 36	442	100.0	379	9	Sequence 671, App
C 37	442	100.0	379	9	Sequence 671, App
C 38	442	100.0	379	9	Sequence 671, App
C 39	442	100.0	379	9	Sequence 671, App
C 40	442	100.0	379	9	Sequence 671, App
C 41	442	100.0	379	9	Sequence 671, App
C 42	442	100.0	379	9	Sequence 671, App
C 43	442	100.0	379	9	Sequence 671, App
C 44	442	100.0	379	9	Sequence 671, App
C 45	442	100.0	379	9	Sequence 671, App
C 46	442	100.0	379	9	Sequence 671, App
C 47	442	100.0	379	9	Sequence 671, App
C 48	442	100.0	379	9	Sequence 671, App
C 49	442	100.0	379	9	Sequence 671, App
C 50	442	100.0	379	9	Sequence 671, App
C 51	442	100.0	379	9	Sequence 671, App
C 52	442	100.0	379	9	Sequence 671, App
C 53	442	100.0	379	9	Sequence 671, App
C 54	442	100.0	379	9	Sequence 671, App
C 55	442	100.0	379	9	Sequence 671, App
C 56	442	100.0	379	9	Sequence 671, App
C 57	442	100.0	379	9	Sequence 671, App
C 58	442	100.0	379	9	Sequence 671, App
C 59	442	100.0	379	9	Sequence 671, App
C 60	442	100.0	379	9	Sequence 671, App
C 61	442	100.0	379	9	Sequence 671, App
C 62	442	100.0	379	9	Sequence 671, App
C 6					


```
; Publication No. US20030069180A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secret, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.527C2
; CURRENT APPLICATION NUMBER: US/10/146.502
; CURRENT FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 671
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 20, 21
; OTHER INFORMATION: n = A,T,C or G
US-10-146-502-671

Alignment Scores:
Pred. No.: 2,81e-46 Length: 379
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x US-10-146-502-671 (1-379)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db 301 ATGAGGCTTCCTAGTCTCTTCCAGCGCTGCTGTATCTCTGCTCTCTCTCCATCTTC 242
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 241 TCCACAGAAGGGAAGAGGCGCTCTGCCAAGGCGCTGTCAGGAGGAGAACAGGCTCTGC 182
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
Db 181 TGCCACCAGTCCCTAGTCCCAACTCAACAACTGAAAGGACATCATGTGAGGCTCTGT 122
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
Db 121 AAACCATGCAAGCTTGAGCCAGAGCCCGCCCTTTGGGTGGTGGCTGGGGGCACTCCACAG 62
Qy 81 Val 81
Db 61 GTG 59

RESULT 5
US-09-981-353-177
; Sequence 177, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981.353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 177
; LENGTH: 800
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1736965CB1
US-09-981-353-177

Alignment Scores:
Pred. No.: 7,13e-46 Length: 800
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x US-09-981-353-177 (1-800)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db 25 ATGAGGCTTCCTAGTCTCTTCCAGCGCTGCTGTATCTCTGCTCTCTCTCCATCTTC 84
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 85 TCCACAGAAGGGAAGAGGCGCTCTGCCAAGGCGCTGTCAGGAGGAGAACAGGCTCTGC 144
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
Db 145 TGCCACCAGTCCCTAGTCCCAACTCAACAACTGAAAGGACATCATGTGAGGCTCTGT 204
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
Db 205 AAACCATGCAAGCTTGAGCCAGAGCCCGCCCTTTGGGTGGTGGCTGGGGGCACTCCACAG 264
Qy 81 Val 81
Db 265 GTG 267

RESULT 6
US-10-028-072-149
; Sequence 149, Application US/10028072
; Publication No. US20030004311A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028,072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
```

[illegible]

RESULT 7

US-10-121-049-149

Sequence 149, Application US/10121049

Publication No. US2003002239A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME

FILE REFERENCE: P3330R1C17

CURRENT APPLICATION NUMBER: US/10/121. 049

```

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-904-149

```

Alignment Scores:	
Pred. No.:	7,18e-46
Score:	442.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	9
Length:	804
Matches:	81
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-724-000A-5 (1-81) x US-10-123-904-149 (1-804)

Qy		1	MetArgLeuValLeuSerSerLeuLeuCysAlleLeuLeuCysPheSerIlePhe	20
Dd		42	ATGAGGTCTTAGTCCCTTTTCAGCGTGCTGTATCTCTGTCTGTCTCATCTTC	101
Qy		21	SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys	40
Dd		102	TCCACAGAGGGAGAGGGGCTCTGCCAAGGCCTGGTCAGCAGGAGAACCAGGCTTCG	161
Qy		41	CystHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys	60
Dd		162	TGCCACCGAGTCCCCTAGCCCCACTCAACAACCTGAAGAGACATCATGTGAGGCTCTGT	221
Qy		61	LysProCysIysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln	80
Dd		222	AACCATGCAAAGCTTGAGCCAGAGCCCCGCTTTGGGTGGTGCTGGGGCACCTCCACAG	281
Qy		81	Val .81 282 GTG 284	

RESULT 9

US-10-140-470-149
; Sequence 149, Application US/10140470
; Publication No. US20030022331A1

GENERAL INFORMATION:

APPLICANT:	Baker, Kevin P.
APPLICANT:	Beresini, Maureen
APPLICANT:	DeForge, Laura
APPLICANT:	Denoyers, Luc
APPLICANT:	Filayoroff, Ellen
APPLICANT:	Gao, Wei-Qiang
APPLICANT:	Gerritsen, Mary E.
APPLICANT:	Goddard, Audrey
APPLICANT:	Godowski, Paul J.
APPLICANT:	Gurney, Austin L.
APPLICANT:	Sherwood, Steven
APPLICANT:	Smith, Victoria
APPLICANT:	Stewart, Timothy A.
APPLICANT:	Tumas, Daniel
APPLICANT:	Watanabe, Colin K
APPLICANT:	Wood, William

```

;
; APPLICANT: Zhang, Zemin
;
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;
; FILE REFERENCE: P3330R1C160
;
; CURRENT APPLICATION NUMBER: US/10/140,470
;
; CURRENT FILING DATE: 2002-05-06
;

```

Prior Application:

Alignment Scores:	
Pred. No.:	7.18e-46
Score:	442.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	9
Length:	804
Matches:	81
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-724-000A-5 (1-81) x US-10-140-470-149 (1-804)

Qy	1	MetArgLeuValLeuSerSerLeuLeuCysileLeuLeuLeuCysPheSerIlePhe	20
Db	42	ATGAGGCTTCGTAGTCCCTTTCCAGCGTGCTGTATCTCTGCTTCTCCTCATCTTC	101
Qy	21	SerThrGluGlyLysArgProAlaLysAlaTrpSerGlyVargThrArgLeuCys	40
Db	102	TCCACAGAGGAAGAGCGCTCTGCCAAGCGCTGGTCAGGCAGGAAACCAGGCTCTGC	161
Qy	41	CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys	60
Db	162	TGCCACCAGTCCCTTAGCCCCAACCTCAACAACCTGAAAGGACATCATGTGAGGCTCTGT	221
Qy	61	LysProCysLysLeuProGluProArgLeuTrpValValProGlyAlaLeuProGln	80
Db	222	AACCATGCAGACTTGAGCCAGAGCCCCCGCTTTGGGTGTGCTGGGGCACCTCCCACAG	281
Qy	81	Val	81
Db	282	GTG	284

RESULT 10

US-10-175-746-149
; Sequence 149, Application US/10175746
; Publication No. US20030027270A1

GENERAL INFORMATION:

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: E3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550

```

; SEO ID NO 149

Alignment Scores:		
Pred. No.:	7,186-46	Length: 804
Score:	442.00	Matches: 81
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	100.00%	Indels: 0
DB:	9	Gaps: 0

US-09-724-000A-5 (1-81) x US-10-175-746-149 (1-804)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
|||
Db 42 ATGAGGGCTTAGTCCTTTCCAGCGTCTGTATCCTGCTCTGCTTCTCATCTTC 101
|||
Qv 21 SerThrGluGlyLvsArqArqProAlaLvsAlaTrpSerGlyArqArqThrArqLeuCys 40

```
Db 102 TCCACAGAGGAAGAGCGTCTCCAGAGCGCTGGTCAGGAGAGAACAGGCTCTGC 161
Qy 41 CysHisArgValProSerProAsnSerThrAenLeuLysGlyHisValArgLeuCys 60
Db 162 TGCCACCGAGTCCCTAGCCCCCACTCAACAACTGAAGGACATCATGTGAGGCTCTGT 221
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 222 AAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 281
Qy 81 Val 81
Db 282 GTG 284

RESULT 11
US-10-176-918-149
; Sequence 149, Application US/10176918
; Publication No. US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Smith, Victoria
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-918-149

Alignment Scores:
Pred. No.: 7,18e-46 Length: 804
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x US-10-176-918-149 (1-804)
Qy 1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuLeuLeuLeuPhe 20
Db 42 ATGAGGCTTCTAGTCTTCCAGCGTGTCTGTATCTCTGCTTCTCTCTCTCTCTCTTC 101
Qy 21 SerThrGluGlyLysArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 102 TCCACAGAGGAAGAGCGTCTCCAGGCGCTGGTCAGGAGAACAGGCTCTGC 161
Qy 41 CysHisArgValProSerProAsnSerThrAenLeuLysGlyHisValArgLeuCys 60
Db 162 TGCCACCGAGTCCCTAGCCCCCACTCAACAACTGAAGGACATCATGTGAGGCTCTGT 221
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 222 AAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 281
Qy 81 Val 81
Db 282 GTG 284

RESULT 13
```

```
Db 222 AAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 281
Qy 81 Val 81
Db 282 GTG 284

RESULT 12
US-10-176-921-149
; Sequence 149, Application US/10176921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C288
; CURRENT APPLICATION NUMBER: US/10/176,921
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-921-149

Alignment Scores:
Pred. No.: 7,18e-46 Length: 804
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x US-10-176-921-149 (1-804)
Qy 1 MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuLeuLeuLeuPhe 20
Db 42 ATGAGGCTTCTAGTCTTCCAGCGTGTCTGTATCTCTGCTTCTCTCTCTCTCTCTTC 101
Qy 21 SerThrGluGlyLysArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 102 TCCACAGAGGAAGAGCGTCTCCAGGCGCTGGTCAGGAGAACAGGCTCTGC 161
Qy 41 CysHisArgValProSerProAsnSerThrAenLeuLysGlyHisValArgLeuCys 60
Db 162 TGCCACCGAGTCCCTAGCCCCCACTCAACAACTGAAGGACATCATGTGAGGCTCTGT 221
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 222 AAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 281
Qy 81 Val 81
Db 282 GTG 284

RESULT 13
```

US-10-137-865-149
; Sequence 149, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-137-865-149

Alignment Scores:
Pred. No.: 7.18e-46 Length: 804
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x US-10-137-865-149 (1-804)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db 42 ATGAGGCTTCCTAGTCTTTCCAGGCTGCTCTGTATCTCTGCTTCTCTCCATCTTC 101
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 102 TCCACAGAAGGGAGAGAGCGTCTGCCAAGGCTGCTCAGGAGGAGAACCCAGGCTCTGC 161
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys 60
Db 162 TGCCACCGAGTCCCTAGCCCCCACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 222 AAACATCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 281
Qy 81 Val 81
Db 282 GTG 284

RESULT 14

US-10-140-474-149
; Sequence 149, Application US/10140474
; Publication No. US20030032156A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C162
; CURRENT APPLICATION NUMBER: US/10140,474
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-474-149

Alignment Scores:
Pred. No.: 7.18e-46 Length: 804
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x US-10-140-474-149 (1-804)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db 42 ATGAGGCTTCCTAGTCTTTCCAGGCTGCTCTGTATCTCTGCTTCTCTCCATCTTC 101
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 102 TCCACAGAAGGAAGAGCGTCTGCCAAGGCTGCTCAGGAGGAGAACCCAGGCTCTGC 161
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys 60
Db 162 TGCCACCGAGTCCCTAGCCCCCACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT 221
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 222 AAACATCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 281
Qy 81 Val 81
Db 282 GTG 284

RESULT 15

US-10-142-431-149
; Sequence 149, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-149

Alignment Scores:
Pred. No.: 7.18e-46 Length: 804
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x US-10-142-431-149 (1-804)

Qy	1	MetArgLeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe	20
Db	42	ATGAGGCTTCTAGTCCCTTCCAGCCCTGCTCTGTATCTCTGCTTCTCTCCATCTTC	101
Qy	21	SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys	40
Db	102	TCCACAGAAGGGAGAGCGTCTCTGCCAAGGCTGTCTCAGCAGGAGAACCCAGGCTCTGC	161
Qy	41	CysHisArgValProSerProAsnSerThrAenLeuLysGlyHisValArgLeuCys	60
Db	162	TGCCACCGAGTCCCTAGCCCACTCAACAAACCTGAAAGGACATCATGTGAGGCTCTGT	221
Qy	61	LysProCysValLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln	80
Db	222	AAACCATGCAAGCTTAGCCAGAGCCCGCCTTTGGGTGGTGGCTGGGGGCACTCCACAG	281
Qy	81	Val	81
Db	282	GTG	284

Search completed: May 17, 2003, 14:37:03
Job time : 152 secs

THIS PAGE BLANK (user)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 03:25:14 ; Search time 159 Seconds
(without alignments)
1147.244 Million cell updates/sec

Title: US-09-724-000A-5

Perfect score: 442

Sequence: 1 MRLVLVSLLCILLCLCSIF.....CKLEPPRLMVVPGALPQV 81

Scoring table: BLOSUM62

-Q/cgn2 1/USPTO.apool/US09724000/runat 16052003 145459 17136/app query.fasta_1.263
-DB-N Genseq_101002 -QFMT=fastap -SURFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-MODE=LOCAL -OUTFMT=ptc -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-USER=US09724000.OCGN 1 1 263 @runat 16052003 145459 17136 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q/cgn2 1/USPTO.apool/US09724000/runat 16052003 145459 17136/app query.fasta_1.263
-DB-N Genseq_101002 -QFMT=fastap -SURFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LIST=45 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-MODE=LOCAL -OUTFMT=ptc -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-USER=US09724000.OCGN 1 1 263 @runat 16052003 145459 17136 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Genseq_101002.*

1: /SID22/cgdata/genseq/genseq-emb1/NA1980.DAT.*
2: /SID22/cgdata/genseq/genseq-emb1/NA1981.DAT.*
3: /SID22/cgdata/genseq/genseq-emb1/NA1982.DAT.*
4: /SID22/cgdata/genseq/genseq-emb1/NA1983.DAT.*
5: /SID22/cgdata/genseq/genseq-emb1/NA1984.DAT.*
6: /SID22/cgdata/genseq/genseq-emb1/NA1985.DAT.*
7: /SID22/cgdata/genseq/genseq-emb1/NA1986.DAT.*
8: /SID22/cgdata/genseq/genseq-emb1/NA1987.DAT.*
9: /SID22/cgdata/genseq/genseq-emb1/NA1988.DAT.*
10: /SID22/cgdata/genseq/genseq-emb1/NA1989.DAT.*
11: /SID22/cgdata/genseq/genseq-emb1/NA1990.DAT.*
12: /SID22/cgdata/genseq/genseq-emb1/NA1991.DAT.*
13: /SID22/cgdata/genseq/genseq-emb1/NA1992.DAT.*
14: /SID22/cgdata/genseq/genseq-emb1/NA1993.DAT.*
15: /SID22/cgdata/genseq/genseq-emb1/NA1994.DAT.*
16: /SID22/cgdata/genseq/genseq-emb1/NA1995.DAT.*
17: /SID22/cgdata/genseq/genseq-emb1/NA1996.DAT.*
18: /SID22/cgdata/genseq/genseq-emb1/NA1997.DAT.*
19: /SID22/cgdata/genseq/genseq-emb1/NA1998.DAT.*
20: /SID22/cgdata/genseq/genseq-emb1/NA1999.DAT.*
21: /SID22/cgdata/genseq/genseq-emb1/NA2000.DAT.*
22: /SID22/cgdata/genseq/genseq-emb1/NA2001A.DAT.*
23: /SID22/cgdata/genseq/genseq-emb1/NA2001B.DAT.*
24: /SID22/cgdata/genseq/genseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	442	100.0	382	24	ABL37082
C 2	442	100.0	485	24	ABL27569
C 3	442	100.0	804	22	AAS21318
C 4	442	100.0	806	24	AAD27025
C 5	442	100.0	908	22	AAF97898
C 6	400	90.5	797	21	AA972230
C 7	400	90.5	801	21	AA972224
C 8	228.5	51.7	744	24	AAD27024
C 9	178	40.3	4159	24	AAD27026
C 10	92	20.8	487	22	AAS39569
C 11	92	20.8	487	22	AAK8908
C 12	91	20.6	619	23	ABK41859
C 13	86.5	19.6	2167	15	AAQ57015
C 14	85	19.2	1622	21	AAQ98948
C 15	85	19.2	1688	17	AAT27617
C 16	85	19.2	1979	17	AAT27616
C 17	85	19.2	2010	24	ABK84201
C 18	85	19.2	2030	15	AAQ63134
C 19	85	19.2	2030	17	AAT18996
C 20	85	19.2	2030	17	AAT30031
C 21	84.5	19.1	2958	20	AAQ76729
C 22	81	18.3	858	21	AAF18164
C 23	80.5	18.2	65140	22	AAD17184
C 24	80.5	18.2	125401	22	AAD17186
C 25	79.5	18.0	1385	23	ABL23395
C 26	79.5	18.0	2206	21	ABL74372
C 27	79.5	18.0	3549	23	ABL23394
C 28	79	17.9	433	14	AAQ59506
C 29	79	17.9	1735	18	AAT77840
C 30	79	17.9	2132	18	AAT77838
C 31	79	17.9	5535	20	AAZ32057
C 32	79	17.9	5535	20	AAZ21355
C 33	79	17.9	5535	22	AAC90314
C 34	78	17.6	1389	23	AAQ84522
C 35	78	17.6	1813	16	AAQ88760
C 36	78	17.6	1898	18	AAQ79634
C 37	78	17.6	2400	23	AAQ92478
C 38	77.5	17.5	7725	22	AAK87157
C 39	77	17.4	3015	24	ABI99621
C 40	77	17.4	17803	22	AAK68676
C 41	77	17.4	34289	22	AAK68677
C 42	77	17.4	34289	22	AAK85168
C 43	76.5	17.3	406	22	AAH42816
C 44	76.5	17.3	1539	24	ABN95921
C 45	76.5	17.3	1539	24	ABL63095

ALIGNMENTS

RESULT 1
ABL37082/c
ID ABL37082 standard; cDNA; 382 BP.
XX ABL37082;
AC ABL37082;
XX
DT 08-APR-2002 (first entry)
XX
DE Human colon tumour antigen polynucleotide SEQ ID NO:671.
XX
KW Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KW colon tumour metastatic antigen; diagnosis; gene; ss.
OS Homo sapiens.
XX
FN WO200196388-A2.
XX
PD 20-DEC-2001.

```

XX PF 08-JUN-2001; 2001WO-US18557.
XX PF
XX PR 09-JUN-2000; 2000US-210899P.
XX PR 20-FEB-2001; 2001US-270216P.
XX PR
XX PA (CORI-) CORIXA CORP.
XX PA
XX PI Jiang Y, Harlocker SL, Secrist H;
XX PI
XX DR WPI; 2002-114514/15.
XX DR
XX PT Novel isolated colon tumor polynucleotide differentially expressed in
XX PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
XX PT useful for inhibiting development of cancer in patient
XX PT
XX PS Claim 1; SEQ ID 671; 105pp; English.
XX PS
XX CC ABL36412 to ABL38645 represent human colon tumor antigen cDNA clones (I)
XX CC which were isolated from human colon tumor and colon metastatic tumor
XX CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
XX CC production. (I) can be used for stimulating and/or expanding T cells
XX CC specific for a tumour protein on contact with the T cells. They are also
XX CC useful for inhibiting the development of cancer in a patient. (I) can be
XX CC used as probes or primers for nucleic acid hybridisation, for preparing
XX CC mutant species primers, or primers for use in genetic constructions. (I)
XX CC can be used in the diagnosis of a colon tumour.
XX CC
XX SQ Sequence 382 BP; 76 A; 94 C; 122 G; 86 T; 4 other;
XX SQ
Alignment Scores:
Pred. No.: 2,26e-38 Length: 382
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-724-000A-5 (1-81) x ABL37082 (1-382)
Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuLeuCysPheSerIlePhe 20
Db 301 ATGAGGCTTCTAGTCTTCCAGCGCTGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 242
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 241 TCCACAGAAGGGAAGAGGCGCTCTGCCAAGGCGCTGTGTCAGGAGAGAACCCAGGCTCTGC 182
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
Db 181 TGCCACCGAGTCCCTAGCCCCCACTCAACAACTGAAGGACATCATGTGAGGCTCTGT 122
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80
Db 121 AAACCATCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGTGGGCGACTCCACAG 62
Qy 81 Val 81
Db 61 GTG 59

RESULT 2
ABK27569
ID ABK27569 standard; cDNA; 485 BP.
XX AC ABK27569;
XX DT 09-APR-2002 (first entry)
XX DE Human colon cancer expressed sequence tag, Seq ID no 6.
XX KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
XX XW expressed sequence tag.

```

```

OS Homo sapiens.
XX WO200196390-A2.
XX 20-DEC-2001.
XX 08-JUN-2001; 2001WO-US18577.
XX 09-JUN-2000; 2000US-210821P.
XX 18-DEC-2000; 2000US-256571P.
XX 10-MAY-2001; 2001US-290240P.
XX (CORI-) CORIXA CORP.
XX Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;
XX WPI; 2002-139708/18.
XX Novel isolated polynucleotide encoding a polypeptide comprising a
XX portion of colon tumour protein, useful for detection, diagnosis and
XX therapy of human colon cancer
XX Claim 1; Page 151-152; 220pp; English.
XX The invention relates to an isolated polynucleotide (I) encoding a
XX polypeptide (II) comprising at least a portion of a colon tumour
XX protein. (I), (II) and antibody (III) to (II) are useful for determining
XX the presence of a cancer in a patient. (I), (II) or antigen presenting
XX cells expressing (I) is useful for stimulating and/or expanding T cells
XX specific for a tumour protein, by contacting T cells with (I), (II) or
XX antigen-presenting cells that express (I), under conditions and for a
XX time sufficient to permit the stimulation and/or expansion of T cells.
XX (I), (II), or antigen presenting cells that express (I) are useful for
XX treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
XX isolated from a patient with (I), (II) or antigen presenting cells that
XX express (II), such that T cells proliferate, and administering to the
XX patient an effective amount of the proliferated T cells, thus inhibiting
XX the development of a cancer in the patient. (I) or (II) is useful in
XX vaccines and pharmaceutical compositions for prevention and treatment
XX of colon malignancies and for the diagnosis and monitoring of such
XX cancers. (I), (II) or (III) is useful for detection, diagnosis and/or
XX therapy of human colon cancer. (I) is useful as a probe or primer for
XX nucleic acid hybridisation, and in the design and preparation of
XX ribozyme molecules for inhibiting expression of (II) in tumour cells.
XX ABK27564-ABK27807 represent novel human colon cancer coding
XX sequences and primers of the invention.
XX SQ Sequence 485 BP; 109 A; 155 C; 124 G; 97 T; 0 other;
XX SQ
Alignment Scores:
Pred. No.: 3,07e-38 Length: 485
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-724-000A-5 (1-81) x ABK27569 (1-485)
Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db 79 ATGAGGCTTCTAGTCTTCCAGCGCTGCTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTTC 138
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 139 TCCACAGAAGGGAAGAGGCGCTCTGCCAAGGCGCTGTGTCAGGAGAGAACCCAGGCTCTGC 198
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60
Db 199 TGCACCGAGTCCCTAGCCCCCACTCAACAACTGAAGGACATCATGTGAGGCTCTGT 258
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGln 80

```


DR WPI; 2000-482912/42.
 XX New isolated polynucleotide useful for diagnosis and/or treatment of
 PT colon cancer and autoimmune disease -
 XX
 PS Claim 3; Page 34; 41pp; English.
 XX
 CC This sequence represents human CASB611 cDNA. This gene exhibits
 CC colon-specific expression and is highly expressed in the rectum.
 CC The invention relates to human CASB cDNA sequences CASB611, CASB500,
 CC CASB501, CASB502, CASB505 and CASB507 (AAA72224-A72229, respectively)
 CC and also to human CASB partial cDNA sequences (AAA72230-A72235) derived
 CC from expressed sequence tags (ESTs). Expression of the human CASB genes
 CC (with the exception of CASB611) is associated with colon tumours, and
 CC the encoded proteins (sequences not given in the specification)
 CC represent colon tumour-associated antigens. The cDNA sequences may be
 CC used in diagnosing the presence or a susceptibility to a disease related
 CC to the presence, expression or activity of CASB genes. Such diseases
 CC include autoimmune diseases and especially colon cancer. The nucleic
 CC acid sequences may also be used in genetic vaccines for the prophylaxis
 CC or therapeutic treatment of colon cancer and autoimmune diseases.
 XX
 SQ Sequence 801 BP; 165 A; 181 C; 255 G; 200 T; 0 other;

Alignment Scores:
 Pred. No.: 1.79e-33 Length: 801
 Score: 400.00 Matches: 78
 Percent Similarity: 96.34% Conservative: 1
 Best Local Similarity: 95.12% Mismatches: 2
 Query Match: 90.50% Indels: 2
 DB: 21 Gaps: 0

US-09-724-000A-5 (1-81) x AAA72224 (1-801)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysLeuLeuLeuLeuCysPheSerIlePhe 20
 Db 769 ATGAGCTTCTAGTCTTCCAGCTGCTCTGTATCTCTGCTTCTCTCATCTTC 710
 Qy 21 SerThrGluGlyLeuArgArgProAlaLys-AlaTrpSerGlyArgThrArgLeuCy 40
 Db 709 TCCACAGAGGGAGAGAGCGCTCTGCCAACCGCTGTGTCAGGAGGAGAACCGAGCTCTG 650
 Qy 40 sCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCy 60
 Db 649 CTGCCACCGAGTCTCTAGCCCCCACTCAACAACTGAAGAGACATCATGTGAGGCTCTG 590
 Qy 60 sLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProGl 80
 Db 589 TAAACCATGCAAGCTTGAGCCAGAGCCCCCGCTTTGGTGTGCTGGGGA-GTCCACACA 531

Qy 80 nVal 81
 Db 530 GGTG 527

RESULT 8

AAD27024

ID AAD27024 standard; DNA; 744 BP.

XX AAD27024;

AC AAD27024;

DT 09-APR-2002 (first entry)

XX Mouse Secreted epithelial colon stromal-1 (Secs-1) gene.
 XX Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnary; antiinfertility;
 KW gynaecological; antitumor; antiinflammatory; cancer; cell therapy;
 KW mouse; ds.
 XX Mus musculus.

XX Key Location/Qualifiers
 FH 38..274
 FT /tag= a
 FT /product= "Mouse Secs-1 protein"
 FT 38..109
 FT sig_peptide
 FT /tag= b
 FT 110..271
 FT mat_peptide
 FT /tag= c
 FT /product= "Mature mouse Secs-1 protein"

WO200198497-A1.

XX 27-DEC-2001.

XX 28-NOV-2000; 2000WO-US32479.

XX 21-JUN-2000; 2000US-0599087.

XX 28-NOV-2000; 2000US-0724000.

XX (AMGE-) AMGEN INC.

XX Polverino AJ, Luethy R;

XX WPI; 2002-122281/16.

XX P-PSDB; AAE16479.

XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 PT useful for diagnosing, treating and preventing hematopoietic disorder,
 PT osteoporosis, Paget's disease, cancer, diabetes -

XX Claim 1; Fig 1; 134pp; English.

XX The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as haematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is mouse Secs-1
 CC gene.

XX Sequence 744 BP; 209 A; 222 C; 155 G; 158 T; 0 other;

SQ Alignment Scores:

Pred. No.: 3.37e-15 Length: 744
 Score: 228.50 Matches: 45
 Percent Similarity: 67.90% Conservative: 10
 Best Local Similarity: 55.56% Mismatches: 23
 Query Match: 51.70% Indels: 3
 DB: 24 Gaps: 1

US-09-724-000A-5 (1-81) x AAD27024 (1-744)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysLeuLeuLeuCysPheSerIlePhe 20

Db 38 ATGAGCTTCTAGCTTTCGGTCTGCTCTGATGCTGCTCTCTCTCTCTCTCTCTCTC 97

Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40

Db 98 TCCTCAGAGGGAGAGACATCTCTGCCAAGTCTCTTGAACCTCAGCGC-----TGC 148

Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCys 60

Db 149 TGTACCTATCTCTCTAGATCCAAAGCTGACAACTGGGAAGGAACCAACACAGGCGCTGC 208

Qy 61 LysProCysLysLeuGluProGluProGluProGluTrpValProGlyAlaLeuProGln 80
 Db 209 AGACTCTGCAGAAACAGCTACCAAGTCAAGTCATGGTGGTGGCTCTCCACAG 268
 Qy 81 Val 81
 Db 269 ATA 271
 RESULT 9
 AAD27026
 ID AAD27026 standard; DNA; 4159 BP.
 XX
 AC AAD27026;
 XX
 DT 09-APR-2002 (first entry)
 XX
 DE Human Secreted epithelial colon stromal-1 (Secs-1) gene.
 XX
 KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; osteoparathic; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
 KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
 KW human; ds.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 PH 1..4159
 FT CDS
 FT /*tag= a
 FT /product= "Human Secs-1 protein"
 FT /note= "CDS does not include stop codon"
 FT /partial
 FT 1..69
 FT /*tag= d
 FT /note= "Encodes Secs-1 peptide (AAE16484)"
 FT /note= 70..2626
 FT /*tag= d
 FT /cons.splices= (5'site:NO, 3'site:NO)
 FT 2627..2725
 FT /*tag= d
 FT /note= "Encodes Secs-1 peptide (AAE16485)"
 FT /note= 2726..4075
 FT /*tag= d
 FT /cons.splices= (5'site:NO, 3'site:NO)
 FT 4076..4159
 FT /*tag= d
 FT /note= "Encodes Secs-1 peptide (AAE16486)"
 XX
 XX WO200198497-A1.
 XX
 XX 27-DEC-2001.
 XX
 XX 28-NOV-2000; 2000WO-US32479.
 XX
 XX 21-JUN-2000; 2000US-0599087.
 XX
 XX 28-NOV-2000; 2000US-0724000.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Polverino AJ, Luethy R;
 XX
 XX WPI: 2002-122281/16.
 XX
 XX P-PSDB: AAE16481, AAE16484, AAE16485, AAE16486.
 XX
 XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 XX useful for diagnosing, treating and preventing hematopoietic disorder,
 XX osteoporosis, Paget's disease, cancer, diabetes -
 XX
 XX Disclosure; Fig 4; 134pp; English.
 XX
 XX The present invention relates to an isolated murine or human secreted

CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as haematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is human Secs-1
 CC gene.
 XX
 SQ Sequence 4159 BP; 1009 A; 1037 C; 983 G; 1110 T; 20 other;
 Alignment Scores:
 Pred. No.: 7.58e-09 Length: 4159
 Score: 178.00 Matches: 37
 Percent Similarity: 62.12% Conservative: 4
 Best Local Similarity: 56.06% Mismatches: 13
 Query Match: 40.27% Indels: 13
 DB: 24 Gaps: 1
 US-09-724-000A-5 (1-81) x AAD27026 (1-4159)
 Qy 24 GlyLysArgProAlaLysAlaTrpSerGlyArgThrArgLeuCysCysHisArg 43
 Db 2635 GGGGAAGAGGGCTCTGCCAAGGGCTGTGTCAGGAGGAGACCCAGGCTCTGCTGCCACCGA 2694
 Qy 44 ValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCysLysProCys 63
 Db 2695 GTCCCTAGCCCCAACTCAACAACCTCAAGGTAAGTACCCCCACCTCGTCAGAC-TGT 2753
 Qy 64 LysLeuGluPro-----GluProArgLeu 71
 Db 2754 GGGCAGAAAGTTTACAGTGGCCATGGGACGACACACACACTGATCAGCCCCACCCCA 2813
 Qy 72 TrpValValProGlyAla 77
 Db 2814 TGGCTGGCATCAGGCTCT 2831
 RESULT 10
 AAS39569
 ID AAS39569 standard; cDNA; 487 BP.
 XX
 AC AAS39569;
 XX
 XX 17-DEC-2001 (first entry)
 XX
 XX cDNA encoding novel human colon associated polypeptide #222.
 XX
 XX Human; colon cancer; congenital abnormality; infection; colitis;
 XX inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
 XX intestinal inflammatory disorder; malabsorption syndrome; gastric;
 XX sigmoid disease; antibacterial; antiviral; antiinflammatory;
 XX cyostatic; ss.
 XX
 XX Homo sapiens.
 XX
 XX WO200155302-A2.
 XX
 XX 02-AUG-2001.
 XX
 XX 17-JAN-2001; 2001WO-US01240.
 XX
 XX 31-JAN-2000; 2000US-0179065.
 XX
 XX 04-FEB-2000; 2000US-0180628.
 XX
 XX 24-FEB-2000; 2000US-0184664.
 XX
 XX 02-MAR-2000; 2000US-0186350.
 XX
 XX 16-MAR-2000; 2000US-0189874.
 XX
 XX 17-MAR-2000; 2000US-0190076.

sequences encoding for them. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of disorders of the colon including colon cancer, congenital abnormalities (e.g. atresia and stenosis), bacterial and viral infections, inflammatory bowel disease (IBD), neoplastic cell disorders (e.g. polyps and adenomas, intestinal inflammatory disorders, colitis, colonic inflammation, diarrhoea and dysentery, malabsorption syndromes (e.g. lactose intolerance), intestinal obstruction and sigmoid diseases. The polynucleotide sequences of the invention can also be used in gene therapy. AAS39348-AAS39581 represent cDNA sequences encoding for the novel human colon associated polypeptides of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 487 BP; 135 A; 156 C; 100 G; 95 T; 1 other;

Alignment Scores:

Pred. No.:	0.741	Length:	487
Score:	92.00	Matches:	16
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	20.81%	Indels:	0
DB:	22	Gaps:	0

US-09-724-000A-5 (1-81) x AAS39569 (1-487)

Oy 38 ArgLeuCySvHisArgValProSerProAsnSerThrAsnLeuLys 53

|||||
Db 3 AGGCTCTGCTGCCACCGAGTCCTTAGCCCACTCACAACCTGAAA 50

RESULT 11

AAK88908

ID AAK88908 standard; cDNA; 487 BP.

XX AAK88908;

XX 05-NOV-2001 (first entry)

XX Human digestive system antigen coding sequence SEQ ID NO: 1224.

XX Human; digestive system antigen; gene therapy; cancer; appendicitis; ulcerative colitis; infection; Hirschsprung's disease; chronic colitis; digestive system disorder; Meckel's diverticulum; ss.

XX Homo sapiens.

XX WO200155314-A2.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001WO-US01324.

XX 31-JAN-2000; 2000US-0179065.

XX 04-FEB-2000; 2000US-0180628.

XX 24-FEB-2000; 2000US-0184664.

XX 02-MAR-2000; 2000US-0186350.

XX 16-MAR-2000; 2000US-0189874.

XX 17-MAR-2000; 2000US-0190076.

XX 18-APR-2000; 2000US-0198123.

XX 19-MAY-2000; 2000US-0205515.

XX 07-JUN-2000; 2000US-0209467.

XX 28-JUN-2000; 2000US-0214886.

XX 30-JUN-2000; 2000US-0215135.

XX 07-JUL-2000; 2000US-0216647.

XX 07-JUL-2000; 2000US-0216880.

XX 11-JUL-2000; 2000US-0217487.

XX 11-JUL-2000; 2000US-0217496.

XX 14-JUL-2000; 2000US-0218290.

XX 26-JUL-2000; 2000US-0220963.

XX 26-JUL-2000; 2000US-0220964.

XX 14-AUG-2000; 2000US-0224518.

XX 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-023935.
PR 13-OCT-2000; 2000US-0239357.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
DR
DR
DR WPI; 2001-502630/55.
DR P-PSDB; AAM93135.
XX
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases.
XX
XX
PS Claim 1; SEQ ID NO 1224; 986bp; English.
XX
XX
CC The present invention provides the protein and coding sequences of a
CC number of human digestive system antigens. These can be used in the
CC diagnosis, treatment and prevention of digestive system disorders,
CC including cancer, Meckel's diverticulum, bacterial or parasitic
CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
CC ulcerative colitis. The present sequence is a cDNA encoding a digestive
CC system antigen of the invention.
XX
XX
SQ Sequence 487 BP; 135 A; 156 C; 100 G; 95 T; 1 other;
Alignment Scores:
Pred. No.: 0.741 Length: 487
Score: 92.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.81% Indels: 0

DB: 22 Gaps: 0
US-09-724-000A-5 (1-81) x AAK88908 (1-487)
Qy 38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeuLys 53
Db 3 AGGCTCTGTCGCCACCGAGTCCTAGCCCACTCAACAAACCTGAAA 50
RESULT 12
ABK41859/C
ID ABK41859 standard; cDNA; 619 BP.
XX
XX AC ABK41859;
XX
XX DT 21-MAY-2002 (first entry)
XX
XX DE cDNA encoding novel human connective tissue related polypeptide #247.
XX
XX KW Human; connective tissue related disorder; cancer; gene therapy;
KW cytostatic; gene; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO20015343-A1.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US01322.
XX
XX PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216847.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231245.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-565190/63.
DR P-PSDB; AAU86681.
XX

Nucleic acid encoding novel connective tissue associated polypeptides, used in diagnosing, preventing, treating or ameliorating a disorder such as cancer or rheumatoid arthritis -

Claim 4; SEQ ID No 257; 673pp; English.

XX The present invention relates to the isolation of novel human connective tissue related polypeptides (AAU86435-AAU86923) and the polynucleotide (cDNA and genomic) sequences encoding them. The sequences of the invention are useful in the diagnosis, treatment, prevention and/or prognosis of diseases associated with connective tissue(s), including cancer. The polynucleotide sequences of the invention are also useful in gene therapy. ABK41613-ABK42101 represent cDNA sequences encoding the novel human connective tissue related polypeptides.
CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 619 BP; 129 A; 194 C; 191 G; 103 T; 2 other;

Alignment Scores:

Pred. No.: 1.29 Length: 619
Score: 91.00 Matches: 24
Percent Similarity: 37.97% Conservative: 6
Best Local Similarity: 30.38% Mismatches: 23
Query Match: 20.59% Indels: 26
DB: 23 Gaps: 3

US-09-724-000A-5 (1-81) x ABK41859 (1-619)

Qy 3 LeuLeuValLeuSerSerLeuLeuCysLeuLeuLeuLeuCysPheSerIlePheSerThr 22
Db 426 CTTCCTGCTCTGCAGTCGCTTCAGTGTCTCCACCTCTGCCTCAGAGCTGGGAGGA 367
Qy 23 GluGlyLysArg-----HisHisValArgLeuLeuCysLys 28
Db 366 GCGCGAGGAATACAGATACATAAAGTGTTCGAGAGCCTGCCCAAACTGAGCTCCC 307
Qy 29 AlaLysAlaTrpSerGlyArgArgThrArgLeuCysCysHisArgValProSerProAsn 48
Db 306 GGGAACGTTGGCAGGCT-----TGCCATCGCGCACCACCAAGGCCCAA 265
Qy 49 SerThrAsnLeuLysGly-----HisHisValArgLeuLeuCysLys 61
Db 264 AACWCTAGCTGCAGGCGCCGCCACCCCTGGCCCATCATCTCTTCCTGGCGGAG 208

RESULT 13
AAQ57015
ID AAQ57015 standard; DNA; 2167 BP.
XX
AC AAQ57015;
XX
DT 31-AUG-1994 (first entry)
XX
DE PKC gamma.
XX
KW 110 kD catalytic subunit; phosphatidyl inositol 3-kinase;
KW transformation; Schizosaccharomyces pombe; nmt promoter; thiamine;
KW PtdIns 3-kinase; assay; detection; cell growth; regulation; cancer;
KW blood vessel plaques; ss.
XX
OS Bos taurus.
XX
XX WO9403609-A.
XX
XX 17-FEB-1994.
XX
XX 05-AUG-1993; 93WO-GB01651.
XX
XX 05-AUG-1992; 92GB-0016654.
XX
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY.
XX
XX Goode NT, Nurse PM, Parker PUJ, Waterfield MD;
XX
XX WPI; 1994-065697/08.
XX
XX Eukaryotic cells transformed with mammalian phospholipid or
XX protein kinase DNA - useful in assays for compounds involved in
XX cell growth regulation and for treating cancers
XX
XX Disclosure; Page 41-42; 71pp; English.
XX
XX The sequences given in AAQ57014-17 encode protein kinase C (PKC)
XX epsilon, gamma, delta and nu respectively. These sequences were
XX transformed into Schizosaccharomyces pombe cells under the regulatory
XX control of the nmt promoter in an embodiment of the invention. In the
XX presence of thiamine the promoter is inactive and the cells carrying
XX the PKC plasmids grow as the parental strain. In the absence
XX of thiamine the nmt promoter functions and the PKC is induced. PKC
XX activity is substantially increased under these conditions. Cells
XX containing constructs such as this, are useful in assays for detecting
XX compounds involved in cell growth regulation. It is also used as the
XX basis for detecting compounds for treating cancers and the formation
XX of blood vessel plaques.
XX
XX Sequence 2167 BP; 428 A; 700 C; 611 G; 428 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 19.5 Length: 2167
Score: 86.50 Matches: 26
Percent Similarity: 39.76% Conservative: 7
Best Local Similarity: 31.33% Mismatches: 19
Query Match: 15.57% Indels: 31
DB: 15 Gaps: 4
US-09-724-000A-5 (1-81) x AAQ57015 (1-2167)
Qy 18 SerIlePheSerThrGluGlyValArg-----ProAlaLysAlaTrp----- 32
Db 1882 TCGAACAGTTCTTCACTCGGCGCGCGCTGACACCCCTGACCGCTGTTCTGG 1941
Qy 33 -----SerGlyArgArgThrArgLeuCysCysHis 42
Db 1942 CCAGCATCGACCGAGTCCAGGCTTCACCTATGTCAACCCGATTTCGTGACC 2001
Qy 43 ArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCysLysPro 62
Db 2002 CGGATGCGCGCAGCCATCATGCCAA-----CGCTGTGCGCAGTCA 2043

Qy 63 Cys-----LysLeuGluProGluProArgLeuTrp 72
Db 2044 TGTAAATCCACCTGCGCCACCGAGGCTCCACCGGCTCCTCTCCGCGCGCTTGG 2103
Qy 73 ValValPro 75
Db 2104 CCCTCGGCT 2112
RESULT 14
AAC98948/c
ID AAC98948 standard; cDNA; 1622 BP.
XX
AC AAC98948;
XX
DT 09-MAR-2001 (first entry)
XX
DE Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:176.
XX
KW Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW detection; diagnosis; identification; cytostatic; neuroprotective;
KW neotropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW linkage analysis; tissue identification; tissue typing; forensic;
KW neural; immune system; muscular; reproductive; gastrointestinal;
KW pulmonary; cardiovascular; renal; proliferative; ss.
XX
OS Homo sapiens.
XX
XX WO200055320-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05989.
XX
XX 12-MAR-1999; 99US-0124270.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI; 2000-579444/54.
XX
XX P-PSDB; AAB54183.
XX
XX New nucleic acid that is a pancreatic cancer antigen for preventing,
XX treating, or ameliorating a medical condition, particular pancreatic
XX cancer, or for use in assays for diagnosing a pathological condition -
XX
XX Claim 1; Page 629; 1379pp; English.
XX
XX AAC98773 to AAC99231 encode the human pancreatic cancer associated
XX proteins, called pancreatic cancer antigens, given in AAB54008 to
XX AAB54466. The human pancreatic cancer antigens have cytostatic,
XX neuroprotective, neotropic, immunomodulatory, relaxant, contraceptive,
XX gynaecological, cardiant and antiinflammatory activities, and can be used
XX in gene therapy. The polynucleotide and proteins can be used for
XX preventing, treating, or ameliorating a medical condition or in assays
XX for diagnosing a pathological condition or a susceptibility to one in a
XX subject. Binding partners to the proteins and the activity of the
XX proteins can be identified. The pancreatic cancer antigens can be used to
XX detect, treat or prevent pancreatic disorders, especially cancer.
XX Agonists and antagonists to the antigens can be screened for. The
XX pancreatic cancer antigen polynucleotides can be used to design nucleic
XX acid hybridisation probes that can be used in chromosome mapping, linkage
XX analysis, tissue identification and/or typing and a variety of forensic
XX and diagnostic methods. The proteins can be used to generate antibodies
XX which are used to purify, detect and target the polypeptides, including
XX both in vivo and in vitro diagnostic and therapeutic methods. The
XX proteins can be used to treat or prevent neural, immune system, muscular,
XX reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
XX proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
XX sequences used in the exemplification of the present invention.
XX

SQ Sequence 1622 BP; 328 A; 486 C; 515 G; 273 T; 20 other;

Alignment Scores:

Pred. No.: 19.4 Length: 1622
Score: 85.00 Matches: 26
Percent Similarity: 47.76% Conservative: 6
Best Local Similarity: 38.81% Mismatches: 17
Query Match: 19.23% Indels: 18
DB: 21 Gaps: 4

US-09-724-000A-5 (1-81) x AAC98948 (1-1622)

Qy 22 ThrGluGlyLysArgArgProAlaLysAla-----TrpSerGlyArgArgThrArgLeu 39
Db 1022 ACTGGGGTCTGCCGCCAGGGCCGAGCGTGACTTTGGCTGTCGGAGAAGGAGCGTT 963
Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisValArg 58
Db 962 TGTTCACCTGCAGTTGGCGCCCACTGCTGGATCATTA----- 921
Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 920 -----GTTCTTGAGCGCTGTTAGCTGGACACCTCGCTTCCCGGAGC 876
Qy 75 ProGlyAlaLeuProGlnVal 81
Db 875 CTGGCTGCCTGCCTCAGATC 855

RESULT 15

AAT27617/C
ID AAT27617 standard; cDNA to mRNA; 1688 BP.

XX AC AAT27617;

XX DT 06-NOV-1996 (first entry)

XX DE Steroid hormone receptor analogue ECDN small mol. variant cDNA.

XX KW Human; foetal lung; steroid hormone; receptor; analogue protein;
KW ECDN protein; cancer; screening; binding molecule; recombinant;
KW identification; anticancer drug; cancerous tissue; primer;
KW probe; antibody; immunohistochemical assay; variant;
KW small molecule; ECDNm; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT CDS. 206..1300
XX FT /*tag= a

XX PN WO9609324-A1.

XX PD 28-MAR-1996.

XX PF 21-SEP-1995; 95WO-JP01909.

XX PR 21-SEP-1994; 94JP-0226270.

XX PA (EISA) EISAI CO LTD.

XX PA (GANK-) ZH GAN KENKYUKAI.

XX PA (CANC-) CANCER INST.

XX PI Nakamura Y, Saito H;

XX DR WPI; 1996-188403/19.

XX DR P-PSDB; AAR96235.

XX PT ECDN protein, a steroid hormone receptor analogue from human foetal
XX lung - is expressed in cancer cells and is useful for cancer
XX diagnosis and drug development

XX PS Claim 3; Pages 25-28; 43pp; Japanese.

XX .XX

CC The present sequence encodes the variant of the human foetal lung
CC derived steroid hormone receptor analogue protein ECDN, designated
CC ECDN small mol. (ECDNm) protein. ECDNm protein is expressed in
CC various cancer cells, therefore screening for ECDNm protein
CC binding mols., using recombinant ECDNm proteins will be useful in
CC the identification of candidate anticancer drugs. Gene expression
CC of ECDNm proteins in cancerous tissues can be studied using
CC primers and probes derived from ECDNm protein cDNA. Antibodies
CC which recognise ECDNm proteins can be used in ECDNm protein
CC immunohistochemical assays.

XX SQ Sequence 1688 BP; 347 A; 562 C; 494 G; 285 T; 0 other;

Alignment Scores:

Pred. No.: 20.4 Length: 1688
Score: 85.00 Matches: 26
Percent Similarity: 47.76% Conservative: 6
Best Local Similarity: 38.81% Mismatches: 17
Query Match: 19.23% Indels: 18
DB: 17 Gaps: 4

US-09-724-000A-5 (1-81) x AAT27617 (1-1688)

Qy 22 ThrGluGlyLysArgArgProAlaLysAla-----TrpSerGlyArgArgThrArgLeu 39
Db 692 ACTGGGGTCTGCCGCCAGGGCCGAGCGTGACTTTGGCTGTCGGAGAAGGAGCGTT 633
Qy 40 CysCysHisArgVal---ProSerProAsnSerThrAsnLeuLysGlyHisValArg 58
Db 632 TGTTCACCTGCAGTTGGCGCCCACTGCTGGATCATTA----- 591
Qy 59 LeuCysLysProCysLysLeuGluPro-----GluProArgLeuTrpValVal 74
Db 590 -----GTTCTTGAGCGCTGTTAGCTGGACACCTCGCTTCCCGGAGC 546
Qy 75 ProGlyAlaLeuProGlnVal 81
Db 545 CTGGCTGCCTGCCTCAGATC 525

Search completed: May 17, 2003, 11:17:01

Job time : 166 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 12:22:51 ; Search time 1740 Seconds
(without alignments)
7502.047 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgagggaataatctgcc.....actcaatgcagacacaaaaa 806

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:

- 1: em_estba:**
- 2: em_esthum:**
- 3: em_estin:**
- 4: em_estmu:**
- 5: em_estov:**
- 6: em_estpl:**
- 7: em_estro:**
- 8: em_htc:**
- 9: gb_est1:**
- 10: gb_est2:**
- 11: gb_htc:**
- 12: gb_est3:**
- 13: gb_est4:**
- 14: gb_est5:**
- 15: em_estfun:**
- 16: em_estom:**
- 17: gb_gss:**
- 18: em_gss_hum:**
- 19: em_gss_inv:**
- 20: em_gss_pln:**
- 21: em_gss_vrt:**
- 22: em_gss_fun:**
- 23: em_gss_mam:**
- 24: em_gss_mus:**
- 25: em_gss_Other:**
- 26: em_gss_pro:**
- 27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486	60.3	557	14	BQ189412 UI-E-EJ1-
C 2	460	57.1	460	9	A1983767 wu20a04.x
C 3	450	55.8	453	10	AW516596 xg01h07.x
C 4	429	53.2	906	12	BE899580 601682443
C 5	410	50.9	543	10	AW970357 EST382438
C 6	406	50.4	413	9	A1339648 qk63a12.x

C 7	398	49.4	450	9	A1948903
C 8	393	48.8	444	9	A1833297
C 9	380	47.1	467	9	A1304380
C 10	379	47.0	430	9	A1283185
C 11	377	46.8	429	9	A1813445
C 12	371	46.0	517	10	AW854263 RC3-CT025
C 13	369	45.8	420	9	AA938765 on44h11.s
C 14	366	45.4	455	9	A1832391
C 15	362	44.9	467	9	A1831407
C 16	359	44.5	405	9	A1833021
C 17	353	43.8	423	9	A1336470
C 18	351	43.5	466	10	AW361500 QV2-CT026
C 19	350	43.4	416	10	AW206923 UI-H-B11-
C 20	346	42.9	431	9	A1832498
C 21	342	42.4	382	9	A1732377
C 22	333	41.3	384	9	A1336626
C 23	323	40.1	331	9	A1246768
C 24	323	40.1	381	9	A1285352
C 25	321	39.8	356	10	AW351839
C 26	320	39.7	381	9	A1732376
C 27	313	38.8	330	10	AW351854
C 28	308	38.2	476	10	AW134688
C 29	307	38.1	460	12	BF001316
C 30	306	38.0	400	9	A1864896
C 31	300	37.2	427	10	AW361498
C 32	288	35.7	441	9	AA587764
C 33	286	35.5	302	12	BF478262
C 34	281	34.9	503	9	AA422178
C 35	277	34.4	484	10	AW970275
C 36	261	32.4	394	9	A1744428
C 37	260	32.3	292	10	AW361503
C 38	259	32.1	357	9	A1766378
C 39	249	30.9	487	9	AA553959
C 40	246	30.5	429	9	AA857922
C 41	230	28.5	415	10	AW361502
C 42	228	28.3	381	9	AA534511
C 43	227	28.2	278	9	AA535314
C 44	218	27.0	385	10	AW591238
C 45	196	24.3	366	9	AA422086

ALIGNMENTS

RESULT 1
BQ189412 557 bp mRNA linear EST 30-APR-2002
LOCUS UI-E-EJ1-ajv-h-16-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
DEFINITION UI-E-EJ1-ajv-h-16-0-UI.r1, mRNA sequence.
ACCESSION BQ189412
VERSION BQ189412.1 GI:20364963
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 557)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa


```

Qy 631 CCAAGACTGCAGAGTCTCTCCATCTTCCAGTCCATTGAGCTCTCTGGCATTTAACTACC 690
Db 160 CCAAGACTGCAGAGTCTCTCCATCTTCCAGTCCATTGAGCTCTCTGGCATTTAACTACC 101
Qy 691 AGCATCCAGTGTCCCAAGGAATCCCTTCTAGGCTCTGACATGAGTCTGCTGGAAG 750
Db 100 AGCATCCAGTGTCCCAAGGAATCCCTTCTAGGCTCTGACATGAGTCTGCTGGAAG 41
Qy 751 AGCATCCAAACAAACAGTAATAATAATAATAATAAATCAATC 790
Db 40 AGCATCCAAACAAACAGTAATAATAATAATAATAAATCAATC 1

RESULT 3
AW516596/c
LOCUS      AW516596      453 bp      mRNA      linear      EST 03-MAR-2000
DEFINITION xq01h07.x1 Soares_NHCEc_cervical_tumor Homo sapiens CDNA clone
ACCESSION IMAGE:2748637 3, mRNA sequence.
VERSION    AW516596.1 GI:7154678
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 453)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL   Unpublished (1997)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -40UP from Gibco
            High quality sequence stop: 451.
FEATURES   Location/Qualifiers
            source          1..453
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:2748637"
                        /clone_lib="Soares_NHCEc_cervical_tumor"
                        /tissue_type="tumor"
                        /lab_host="DH10B (phage-resistant)"
                        /note="Organ: cervix; Vector: pT7T3D-Pac (Pharmacia) with
                        a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                        strand cDNA was primed with a Not I - oligo(dT) primer [5'-
                        TGTTACCAATCTGAAGTGGAGCGCGCCGAGTTTTTTTTTTTTTTTTTTTTTTT
                        T 3']; double-stranded cDNA was ligated to Eco RI
                        adaptors (Pharmacia), digested with Not I and cloned into
                        the Not I and Eco RI sites of the modified pT7T3 vector.
                        Library is normalized; constructed by Bento Soares and
                        M.Fatima Bonaldo."
BASE COUNT      95 a      92 c      142 g      124 t
ORIGIN
Query Match      55.8%; Score 450; DB 10; Length 453;
Best Local Similarity 100.0%; Pred. No. 3.5e-214;
Matches 450; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 343 CTGCTCTCCCTTTTACAGCTTCACAGCAGTGCATGTTGGAGGCTTCATCTCGGG 402
Db 453 CTGCTCTCCCTTTTACAGCTTCACAGCAGTGCATGTTGGAGGCTTCATCTCGGG 394
Qy 403 CTGCAAGACCTCGGAAGTTCAGAACTCCACGCTCTTGTCTCAATTGTCCTCAAC 462
Db 393 CTGCAAGACCTCGGAAGTTCAGAACTCCACGCTCTTGTCTCAATTGTCCTCAAC 334
Qy 463 TTTTCAGAGTATCATGAGCCACCTCACCCACAGGGCTCAGTCGCCACCATGTGGCC 522
Db 333 TTTTCAGAGTATCATGAGCCACCTCACCCACAGGGCTCAGTCGCCACCATGTGGCC 274

```

```

Qy 523 TCTCAGTGCAAACCCAGGAGCATTTCCACATGATACCGGTACAGCTACAAATCCAGAGAC 582
Db 273 TCTCAGTGCAAACCCAGGAGCATTTCCACATGATACCGGTACAGCTACAAATCCAGAGAC 214
Qy 583 CATCAATCCTCTAGAGTGCAGGGTGGCAAGCACCACAGGGTGGTGAACCAAGACTGCAG 642
Db 213 CATCAATCCTCTAGAGTGCAGGGTGGCAAGCACCACAGGGTGGTGAACCAAGACTGCAG 154
Qy 643 AGTCTCTCTCCATCTTCCAGGTCCATTGAGCTCTCTGGCATTTAACTTACCAGCATCCAGTGG 702
Db 153 AGTCTCTCTCCATCTTCCAGGTCCATTGAGCTCTCTGGCATTTAACTTACCAGCATCCAGTGG 94
Qy 703 TCCCAAGGAATCCCTTCTAGGCTCTGACATGAGTCTGCTGGAAGAGCATCCAAACA 762
Db 93 TCCCAAGGAATCCCTTCTAGGCTCTGACATGAGTCTGCTGGAAGAGCATCCAAACA 34
Qy 763 AACAAAGTAATAATAATAATAATAATAAATAACTCAA 792
Db 33 AACAAAGTAATAATAATAATAATAATAAATAACTCAA 4

RESULT 4
BE899580
LOCUS      BE899580      906 bp      mRNA      linear      EST 29-SEP-2000
DEFINITION 50162443F1 NIH_MGC_9 Homo sapiens CDNA clone IMAGE:3952523 5',
ACCESSION mRNA sequence.
VERSION    BE899580
KEYWORDS   BE899580.1 GI:10367234
SOURCE     EST.
            human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 906)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTF
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCMB22 row: f column: 12
            High quality sequence start: 23
            High quality sequence stop: 775.
FEATURES   Location/Qualifiers
            source          1..906
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:3952523"
                        /clone_lib="NIH_MGC_9"
                        /lab_host="DH10B (phage-resistant)"
                        /tissue_type="adenocarcinoma cell line"
                        /note="Organ: ovary; Vector: pOT87; Site_1: XhoI; Site_2:
                        EcoRI; CDNA made by oligo-dT priming. Directionally
                        cloned into EcoRI/XhoI sites using the following 5'
                        adaptor: GGCACGAG(G). Size-selected >500bp for average
                        insert size 1.8kb. Library constructed by Ling Hong in
                        the laboratory of Gerald M. Rubin (University of
                        California, Berkeley) using ZAP-cDNA synthesis kit
                        (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      219 a      282 c      219 g      186 t
ORIGIN
Query Match      53.2%; Score 429; DB 12; Length 906;
Best Local Similarity 100.0%; Pred. No. 1.3e-203;
Matches 429; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 TCTGTATCTCTCTCTGCTTCTTCCATCTTCTCCACAGAGGAGGCGTCTCGCA 116

```


ORIGIN

Query Match 50.4%; Score 406; DB 9; Length 413;
Best Local Similarity 100.0%; Pred. No. 3.7e-192;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 387 AGGGTTTCATCTCGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGTCCTTGCT 446
Db |||||
QY 413 AGGGTTTCATCTCGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGTCCTTGCT 354
Db |||||
QY 447 CAATTGTGCATCAATTTTCAGAGTATCATGAGCCCACTCACCCACAGGCGCTCAGT 506
Db |||||
QY 507 CGCCACCATGTGGGCTCTCCAGTGCACACCGAGCATTCACCATGACCGGTGCAG 566
Db |||||
QY 567 CTACAAATCCAGAGACCATCAATCTGTAGAGTGAGGTGGCAAGCACCCAGGGTGG 626
Db |||||
QY 627 CTGACCAAGACTGCAGAGTCTCTCATCTTCAGGTCCATTTCAGGCTCTGSCATTTAAC 686
Db |||||
QY 687 TACCAAGCATCCAGTGGTCCCCAAGGAATCCCTTCTAGCTCTGACATGAGTCTGCTGG 746
Db |||||
QY 747 AAGAGCATCCAAACAAAGTAATAATAATAATAATACTCAA 792
Db |||||

RESULT 7
AI948903/c
LOCUS wq17c02.x1 NCI_CGAP_Kid12 Homo sapiens cdna clone IMAGE:2471522 3',
DEFINITION mRNA sequence.
ACCESSION AI948903
VERSION AI948903.1 GI:5741213
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 450)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 607 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 438.
Location/Qualifiers
1. .450
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2471522"
/clone_lib="NCI_CGAP_Kid12"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"

FEATURES
source

/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid5 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(Clones IDs 132912-1325831, 1471368-1472903 and
1492104-1493255). Subtraction by Bento Soares and M.
Fatima Bonaldo."

BASE COUNT 94 a 92 c 141 g 123 t
ORIGIN

Query Match 49.4%; Score 398; DB 9; Length 450;
Best Local Similarity 99.8%; Pred. No. 3.8e-188;
Matches 448; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 344 TGTCTCCCTTTACGCTTTCACAGAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGC 403
Db |||||
QY 404 TGCAGGACCTGGGAAAGTTCCAGAACTCCAGCTCCCTTGCTCAATTGTGCCATCAACT 463
Db |||||
QY 390 TGCAGGACCTGGGAAAGTTCCAGAACTCCAGCTCCCTTGCTCAATTGTGCCATCAACT 331
QY 464 TTCAGAGCTATCATGAGCCAACTCACCCACAGGGCTCAGTCGCCACCATGTGGGGCT 523
Db |||||
QY 524 CTCAGTGCAAAACACCGAGCATTTCCACCATGACCGGTCAAGCTACAAATCCAGAGACC 583
Db |||||
QY 584 ATCAATCTCTGCTAGAGTGCAGGGTGGCAAGCACCCAGGGTGGTGCACCAAGACTGCAGA 643
Db |||||
QY 644 GTCTCTCCATCTTTCAGGTCCATTCAGGCTCTCTGGCATTTAACTACCAAGCATCCAGTGGT 703
Db |||||
QY 704 CCCAAGGAATCCCTTCTAGCTCTGACATGATGCTGTGGAAAGAGCATCCAAACAA 763
Db |||||
QY 764 ACAAGTAAATAATAATAATAATACTCAA 792
Db |||||

RESULT 8
AI833297/c
LOCUS at67a07.x1 Barstead colon HPLR87 Homo sapiens cdna clone
DEFINITION IMAGE:2377044 3', mRNA sequence.
ACCESSION AI833297
VERSION AI833297.1 GI:5455277
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 444)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin,
J., Moore, B., Schellberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -40UP from Gibco.

FEATURES

source

1. .444

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2377044"

/clone_lib="Barstead colon HPLRB7"

/sex="male"

/dev_stage="adult, age 25"

/lab_host="DH10B (phage resistant)"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCTAGTAT 3' and 5' ATTACTAGT 3'], digested

with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library constructed by Bob

Barstead."

BASE COUNT 93 a 92 c 141 g 118 t

ORIGIN

Query Match 48.8%; Score 393; DB 9; Length 444;

Best Local Similarity 99.8%; Pred. No. 1.2e-185;

Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 347 CTCCTCCCTTTACGCTTACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGC 406

Db 444 CTCCTCCCTTTACGCTTACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGC 385

Qy 407 AAGGACCTGGGAAAGTTCAGAACTCAGCTGCTTGTCTCAATTGTGCCATCACTTTC 466

Db 384 AAGGACCTGGGAAAGTTCAGAACTCAGCTGCTTGTCTCAATTGTGCCATCACTTTC 325

Qy 467 AGAGCTATCATGAGCAACCTCACCCACAGGCGCTCAGTCGCCACCATGTGGGCTCTC 526

Db 324 AGAGCTATCATGAGCAACCTCACCCACAGGCGCTCAGTCGCCACCATGTGGGCTCTC 265

Qy 527 CAGTGCACCAACCGAGCATTCACCATGACCGGTTCACAGCTTACAAATCCAGACCATC 586

Db 264 CAGTGCACCAACCGAGCATTCACCATGACCGGTTCACAGCTTACAAATCCAGACCATC 205

Qy 587 AATCTGCTAGAGTCAGGGTGGCAAGCACCAGGGTGGCTGACCAAGACTGCAGAGTC 646

Db 204 AATCTGCTAGAGTCAGGGTGGCAAGCACCAGGGTGGCTGACCAAGACTGCAGAGTC 145

Qy 647 TCCTCCATCTTCAGTCCATTCAGCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCC 706

Db 144 TCCTCCATCTTCAGTCCATTCAGCTCCTGGCATTTAACTACCAGCATCCAGTGGTCCC 85

Qy 707 CAAGGAATCCCTTCTAGCCTCTGACATGATGCTGTGGAAAGAGCATCAAAACAACA 766

Db 84 CAAGGAATCCCTTCTAGCCTCTGACATGATGCTGTGGAAAGAGCATCAAAACAACA 25

Qy 767 AGTAATAAATAAATAAATAACTC 790

Db 24 AGTAATAAATAAATAAATAACTC 1

RESULT 9

AI304380/c

LOCUS

DEFINITION q059c12.x1 NCI CGAP Co8 Homo sapiens cDNA clone IMAGE:1912822 3',

mRNA sequence.

ACCESSION AI304380

VERSION AI304380.1 GI:3988069

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 467)

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgaops-r@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.

Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

www-bio.llnl.gov/bbrp/image/image.html

Insert length: 604 Std Error: 0.00

Seq primer: -40UP from Gibco

High quality sequence stop: 438.

Location/Qualifiers

source

1. .467

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:1912822"

/clone_lib="NCI CGAP Co8"

/tissue_type="adenocarcinoma"

/lab_host="DH10B"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from

colon adenocarcinoma, and was then primed with a Not I -

oligo(dT) primer. Double-stranded cDNA was ligated to Eco

RI adaptors (Pharmacia), digested with Not I and cloned

into the Not I and Eco RI sites of the modified pT73

vector. Library is normalized. Library was constructed by

Bento Soares and M. Fatima Bonaldo."

BASE COUNT 96 a 97 c 148 g 126 t

ORIGIN

Query Match 47.1%; Score 380; DB 9; Length 467;

Best Local Similarity 99.8%; Pred. No. 4e-179;

Matches 430; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 325 GAGGTACCCAGCAGCCTCTCTCCCTTTTCAGCCTTCACAGCAGTGAGCTGCAATGTT 384

Db 467 GAGGTACCCAGCAGCCTCTCTCTCCCTTTTCAGCCTTCACAGCAGTGAGCTGCAATGTT 408

Qy 385 GGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCAGAACTCCACGTCCTTGT 444

Db 407 GGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCAGAACTCCACGTCCTTGT 348

Qy 445 CTCATTTGTGCCATCAACTTTCAGAGCTATCATGAGCAACCTCACCCACAGGGCTCA 504

Db 347 CTCATTTGTGCCATCAACTTTCAGAGCTATCATGAGCAACCTCACCCACAGGGCTCA 288

Qy 505 GTGCCACCATGTGGGCTCTCCAGTCAAAACCAACGAGCATTCACCATGACCGGTAC 564

Db 287 GTGCCACCATGTGGGCTCTCCAGTCAAAACCAACGAGCATTCACCATGACCGGTAC 228

Qy 565 AGCTACAAATCCAGAGACCATCAATCTCTGTAGAGTGCAGGGTGGCAAGCACCAGGGT 624

Db 227 AGCTACAAATCCAGAGACCATCAATCTCTGTAGAGTGCAGGGTGGCAAGCACCAGGGT 168

Qy 625 GGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGTCCATTCAGCTCCTCTGGCATTTA 684

Db 167 GGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGTCCATTCAGCTCCTCTGGCATTTA 108

Qy 685 ACTACAGCAGTCCAGTGGTCCCCAAGGAATCCCTTCTAGCTCCTGACATGAGTCTGCT 744

Db 107 ACTACAGCAGTCCAGTGGTCCCCAAGGAATCCCTTCTAGCTCCTGACATGAGTCTGCT 48

Qy 745 GGAAAGAGCAT 755

|||||

```

Db 47 GGAAGAGCAT 37
RESULT 10
A1283185/c
LOCUS
DEFINITION
  qk49g09.x1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1872352 3',
  mRNA sequence.
ACCESSION
  A1283185
VERSION
  A1283185.1 GI:3921418
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 430)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 933 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 403.
  Location/Qualifiers
    1..430
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:1872352"
    /clone_lib="NCI_CGAP C08"
    /tissue_type="adenocarcinoma"
    /lab_hosts="DH10B"
    /notes="Organ: colon; Vector: p7T73D-Pac (Pharmacia) with a
    modified polylinker; 1st strand cDNA was prepared from
    colon adenocarcinoma, and was then primed with a Not I -
    oligo(dT) primer. Double-stranded cDNA was ligated to Eco
    RI adaptors (Pharmacia), digested with Not I and cloned
    into the Not I and Eco RI sites of the modified p7T73
    vector. Library is normalized. Library was constructed by
    Bento Soares and M. Fatima Bonaldo."
  BASE COUNT      88 a 90 c 133 g 119 t
  ORIGIN
    Query Match      47.0%; Score 379; DB 9; Length 430;
    Best Local Similarity 99.8%; Pred. No. 1.2e-178;
    Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 361 TTCACGAGTGAGCTGCAATGTTGGAGGCTTCATCTCGGCTCGAAGGACCCCTGGAA 420
Db 430 TTCACGAGTGAGCTGCAATGTTGGAGGCTTCATCTCGGCTCGAAGGACCCCTGGAA 371
Qy 421 AGTTCAGAACTCCACGCTCTTGTCTCAATTTGTGCCATCAACTTTTCAGAGCTATCATGAG 480
Db 370 AGTTCAGAACTCCACGCTCTTGTCTCAATTTGTGCCATCAACTTTTCAGAGCTATCATGAG 311
Qy 481 CCAACCTCACCCGAGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCACCAACACC 540
Db 310 CCAACCTCACCCGAGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCACCAACACC 251
Qy 541 GAGCATTCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGT 600
Db 250 GAGCATTCACCATGACCGGTCACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGT 191
Qy 601 CGAGGTGGCAAGCAACCAAGGCTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAG 660

Db 190 GCAGGGAGGCAAGCACCAGGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAG 131
Qy 661 GTCCATTTCAGCTCTCTGCGATTAACTACCAAGCATCAGTGGTCCCCAAGGAATCCCTTC 720
Db 130 GTCCATTTCAGCTCTCTGCGATTAACTACCAAGCATCAGTGGTCCCCAAGGAATCCCTTC 71
Qy 721 CTAGCCCTCTGACATGATGCTGTGGAAAGAGCATCCAAACAAACAAGTAATAATAAAT 780
Db 70 CTAGCCCTCTGACATGATGCTGTGGAAAGAGCATCCAAACAAACAAGTAATAATAAAT 11
Qy 781 AATAAATCT 790
Db 10 AATAAATCT 1

RESULT 11
A1813445/c
LOCUS
DEFINITION
  wj06e01.x1 NCI_CGAP_Kid12 Homo sapiens cDNA clone IMAGE:2402040 3',
  mRNA sequence.
ACCESSION
  A1813445
VERSION
  A1813445.1 GI:5424660
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 429)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-remail.nih.gov
  Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
  Emmert-Buck, M.D., Ph.D.
  cDNA Library Preparation: M. Bento Soares, Ph.D.
  cDNA Library Arrayed by: Greg Lennon, Ph.D.
  DNA Sequencing by: Washington University Genome Sequencing Center
  Clone distribution: NCI-CGAP clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  www-bio.llnl.gov/bbrp/image/image.html
  Insert Length: 616 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 389.
  Location/Qualifiers
    1..429
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:2402040"
    /clone_lib="NCI_CGAP Kid12"
    /tissue_type="2_pooled tumors (clear cell type)"
    /lab_hosts="DH10B"
    /notes="Organ: kidney; Vector: p7T73D-Pac (Pharmacia) with
    a modified polylinker; Site 1: Not I; Site 2: Eco RI;
    Plasmid DNA from the normalized library NCI_CGAP Kid5 was
    prepared and ss circles were made in vitro. Following HAP
    purification, this DNA was used as tracer in a subtractive
    hybridization reaction. The driver was PCR-amplified cDNAs
    from a pool of 5,000 clones made from the same library
    (cloneIDs 1323912-1325831, 1471368-1472903 and
    1492104-1493255). Subtraction by Bento Soares and M.
    Fatima Bonaldo."
  BASE COUNT      86 a 90 c 132 g 121 t
  ORIGIN
    Query Match      46.8%; Score 377; DB 9; Length 429;
    Best Local Similarity 99.8%; Pred. No. 1.2e-177;
    Matches 427; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 365 CAGCAGTGAGCTGCAATGTTGGAGGCTTCATCTCGGCTCGAAGGACCCCTGGAAAGTT 424

```

Db 429 CAGCAGTGGAGCTGCAATGTTGGAGGGCTTCATCTCGGCTGCAAGGACCCCTGGGAAAGTT 370
Qy CCAGACTCCAGCTCCTGCTGCTCAATGTCATCACTTTTCAGAGCTATCATGAGCAA 484
Db 369 CCAGACTCCAGCTCCTGCTGCTCAATGTCATCACTTTTCAGAGCTATCATGAGCAA 310
Qy CCTCACCCTCCAGGGCTCAGTCGCGCCACATGTCGGGCTCTCCAGTGCACACCCAGGC 544
Db 309 CCTCACCCTCCAGGGCTCAGTCGCGCCACATGTCGGGCTCTCCAGTGCACACCCAGGC 250
Qy ATTCACCATGACCGCTCAGTGCATCAATCCAGAGACCATCAATCTCTGCTAGAGTCAG 604
Db 249 ATTCACCATGACCGCTCAGTGCATCAATCCAGAGACCATCAATCTCTGCTAGAGTCAG 190
Qy GGTGCGAGCACCACCAAGGCTGACCAAGCTGTCAGAGTCTCTCCATCTTCAGGTCC 664
Db 189 GGAGCGACGACCCCAAGGCTGACCAAGCTGTCAGAGTCTCTCCATCTTCAGGTCC 130
Qy ATTCAGCTCTCGGCTATTAATACAGCATCCAGTGGTCCCAAGGAATCCCTTCCTAG 724
Db 129 ATTCAGCTCTCGGCTATTAATACAGCATCCAGTGGTCCCAAGGAATCCCTTCCTAG 70
Qy CCTCTGACATGAGTCTGCTGGAGAGCAGTCCAAACAAAGTAATAATAATAAT 784
Db 69 CCTCTGACATGAGTCTGCTGGAGAGCAGTCCAAACAAAGTAATAATAATAAT 10
Qy 785 AACTCAA 792
Db 9 AACTCAA 2

RESULT 12
AW854263/c
LOCUS RC3-CT0254-100500-211-g12 CT0254 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW854263
VERSION AW854263.1 GI:7949956
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t=RC3-CT0254-100
500-211-g12&t3=2000-05-10&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 6
High quality sequence stop: 517.
Location/Qualifiers

FEATURES
source
1..517
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CT0254"

/dev stage="Adult"
/note="Organ: colon; Vector: puc18; Site: 1; SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 106 a 118 c 172 g 121 t
ORIGIN

Query Match 46.0%; Score 371; DB 10; Length 517;
Best Local Similarity 100.0%; Pred. No. 1.3e-174;
Matches 371; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 331 CCAGAGCCTCTGCTCTCCCTTTTCAGCTTTCAGAGCTGAGCTCAATGTTGGAGGG 390
Db 434 CCAGAGCCTCTGCTCTCCCTTTTCAGCTTTCAGAGCTGAGCTCAATGTTGGAGGG 375
Qy 391 CTTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGCTCTGTTCTCAAT 450
Db 374 CTTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGCTCTGTTCTCAAT 315
Qy 451 TGTGCCATCAACTTTTCAGAGCTATCATGAGCCAACTTCACCCACAGGGCTCAGTCGCC 510
Db 314 TGTGCCATCAACTTTTCAGAGCTATCATGAGCCAACTTCACCCACAGGGCTCAGTCGCC 255
Qy 511 ACCATGTGGGCTCTCCAGTGCACACCCAGGATTCACCATGAGCGGTTCACAGTAC 570
Db 254 ACCATGTGGGCTCTCCAGTGCACACCCAGGATTCACCATGAGCGGTTCACAGTAC 195
Qy 571 AATTCACAGAGACCATCAATCTCTGCTAGAGTGGAGGGTGGCAAGCACCAGGGTGGCTGA 630
Db 194 AATTCACAGAGACCATCAATCTCTGCTAGAGTGGAGGGTGGCAAGCACCAGGGTGGCTGA 135
Qy 631 CCAAGACTGCAGAGTCTCTCCATCTTCAGTTCATTCAGCTCCTCGCATTTTAATACC 690
Db 134 CCAAGACTGCAGAGTCTCTCCATCTTCAGTTCATTCAGCTCCTCGCATTTTAATACC 75
Qy 691 AGCATCCAGTG 701
Db 74 AGCATCCAGTG 64

RESULT 13
AA938765/c
LOCUS on44h11.s1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1559589 3',
DEFINITION mRNA sequence.
ACCESSION AA938765
VERSION AA938765.1 GI:3096793
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 420)
TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 613 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham

High quality sequence stop: 403.
Location/Qualifiers
1. 420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1559589"
/clone_lib="NCI CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
86 a 88 C 130 G 116 C

BASE COUNT 86 a 88 C 130 G 116 C

FEATURES
source
1. 420
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1559589"
/clone_lib="NCI CGAP_C08"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
86 a 88 C 130 G 116 C

Query Match 45.4%; Score 369; DB 9; Length 420;
Best Local Similarity 99.8%; Pred. No. 1.2e-173;
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 371 TGAGCTGCAATGTTGGAGGGCTTCACTCGGGCTGCAAGACCCCTGGGAAAGTTCAGAA 430
Db 420 TGAGCTGCAATGTTGGAGGGCTTCACTCGGGCTGCAAGACCCCTGGGAAAGTTCAGAA 361

Qy 431 CTCACGTCCTTGTCTCAATGTGGCCATCACTTTCAGAGCTATCATGAGCCACCTCAC 490
Db 360 CTCACGTCCTTGTCTCAATGTGGCCATCACTTTCAGAGCTATCATGAGCCACCTCAC 301

Qy 491 CCCACAGGGCTCAGTCGCGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGCAGATTCCA 550
Db 300 CCCACAGGGCTCAGTCGCGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGCAGATTCCA 241

Qy 551 CCATGACCGGTACAGCTACAAATCCAGAGCCATCAATCTCTGCTAGAGTCGAGGGTGGC 610
Db 240 CCATGACCGGTACAGCTACAAATCCAGAGCCATCAATCTCTGCTAGAGTCGAGGGTGGC 181

Qy 611 AAGCACCAAGGGTGGCTGACCAAGCTGCGAGGTCCTCCATCTTCAGTCCATTCCAG 670
Db 180 AAGCACCAAGGGTGGCTGACCAAGCTGCGAGGTCCTCCATCTTCAGTCCATTCCAG 121

Qy 671 CCTCTGGCATTTAACTACAGCATCCAGTGTGTCCTCCCAAGGAATCCCTTCTAGGCTCCT 730
Db 120 CCTCTGGCATTTAACTACAGCATCCAGTGTGTCCTCCCAAGGAATCCCTTCTAGGCTCCT 61

Qy 731 GACATGATCTGCTGGAAAGAGCATCCAAACAAACAGTAATAATAATAATAATAAATC 790
Db 60 GACATGATCTGCTGGAAAGAGCATCCAAACAAACAGTAATAATAATAATAATAAATC 1

RESULT 14
A1832391/c
LOCUS
DEFINITION
IMAGE:2377181.3', mRNA sequence.
A1832391
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 455)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 420.
Location/Qualifiers
1. 455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2377181"
/clone_lib="Barstead colon HPLRB7"
/sex="male"
/dev_stage="adult, age 25"
/lab_host="DH10B (phage resistant)"
/notes="Organ: colon; Vector: p7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT TGTACGAATCTGAAGTGGAGCGCGCCCTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [5' AATTCTACTAGTAAT 3' and 5' ATTACTAGT 3'], digested with Not I and cloned into the Not I and Eco RI sites of the modified p7T3 vector. Library constructed by Bob Barstead."
97 a 94 C 145 G 119 T

BASE COUNT 97 a 94 C 145 G 119 T

ORIGIN

Query Match 45.4%; Score 366; DB 9; Length 455;
Best Local Similarity 99.8%; Pred. No. 4e-172;
Matches 416; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 341 TCCTGTCTCCCTTTCAGCTTCACAGCAGTGAGTCCGAATGTTGGAGGGCTTCATCTCG 400
Db 455 TCCTGTCTCCCTTTCAGCTTCACAGCAGTGAGTCCGAATGTTGGAGGGCTTCATCTCG 396

Qy 401 GGCTGCAAGGACCTCGGAAAGTTCCAGAACTCCACAGTCTCTGTCTCAATGTGCGCATCA 460
Db 395 GGCTGCAAGGACCTCGGAAAGTTCCAGAACTCCGCGTCTCTGTCTCAATGTGCGCATCA 336

Qy 461 ACTTTCAGAGTATCATGAGCAACCTCACCCACAGGGCTCAGTCCGCCACCATGTGGG 520
Db 335 ACTTTCAGAGTATCATGAGCAACCTCACCCACAGGGCTCAGTCCGCCACCATGTGGG 276

Qy 521 CCTCTCAGTGCAACACCCAGCAGCATTCACCATGACGGTGCAGAGTACAAATCCAGAG 580
Db 275 CCTCTCAGTGCAACACCCAGCAGCATTCACCATGACGGTGCAGAGTACAAATCCAGAG 216

Qy 581 ACCATCAATCTGCTAGAGTGCAGGGTGGCAAGCACCAAGGGTGGCTGACCAAGACTGC 640
Db 215 ACCATCAATCTGCTAGAGTGCAGGGTGGCAAGCACCAAGGGTGGCTGACCAAGACTGC 156

Qy 641 AGAGTCTCTCTCATCTTTCAGTCCATTCAGCTCTCTGCGCATTTAACTACCAAGCTCCAGT 700
Db 155 AGAGTCTCTCTCATCTTTCAGTCCATTCAGCTCTCTGCGCATTTAACTACCAAGCTCCAGT 96

Qy 701 GGTCCTCCAGGAATCCCTTCTAGCTCTGAGCATGATGATGCTGCGAAAGAGCATCC 757
Db 95 GGTCCTCCAGGAATCCCTTCTAGCTCTGAGCATGATGATGCTGCGAAAGAGCATCC 39

RESULT 15
A1831407/c
LOCUS
DEFINITION
IMAGE:2377181.3', mRNA sequence.
A1831407
VERSION
KEYWORDS
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 455)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 467)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: CGAPbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.lnl.gov/bbrp/image/image.html
Insert Length: 756 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 458.

FEATURES
source

```

1. 457
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2407582"
/clone_lib="NCI_CGAP_Lu19"
/tissue_type="squamous cell carcinoma, poorly
differentiated (4 pooled tumors, including primary and
metastatic)"
/dev_stage="adult"
/lab_host="DH10B (phage-resistant)"
/notes="Organ. lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
pooled lung tumor tissue, and was then primed with a Not I
- oligo (dT) primer. Double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT7T3 vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.
Fatima Bonaldo."
94 C 146 Q 132 T

```

BASE COUNT	95 a	94 c	146 g	132 t
ORIGIN				

	Query Match	44.9%	Score 362;	DB 9;	Length 467;
	Best Local Similarity	99.6%;	Pred. No. 4.1e-170;		
	Matches 462;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;
QY	343	CTGTCCTCCCTTTTCAGCCCTTCACAGCAGTGCAGTGCCAATGTTGGAGGGCTTCATCTCGGG	402		
DB	467	CTGTCCTCCCTTTTCAGCCCTTCACAGCAGTGCAGTGCCAATGTTGGAGGGCTTCATCTCGGG	408		
QY	403	CTGCAAGGACCTTGGGAAAGTTTCAGAACTCCACGCTCTTGTCTCTCAATTGTGCCATCAAC	462		
DB	407	CTGCAAGGACCTTGGGAAAGTTTCAGAACTCCACGCTCTTGTCTCTCAATTGTGCCATCAAC	348		
QY	463	TTTCAGAGCTATCATGAGCCAACTCCACCCACAGGGCCTCAGTCGCCACCAATGTGGGCC	522		
DB	347	TTTCAGAGCTATCATGAGCCAACTCCACCCACAGGGCCTCAGTCGCCACCAATGTGGGCC	288		
QY	523	TCTCCAGTGCMAACCCACGGAGCATTTCCACCATGACCCGGTCCACAGCTCAAAATCCAGAGAC	582		
DB	287	TCTCCAGTGCMAACCCACGGAGCATTTCCACCATGACCCGGTCCACAGCTCAAAATCCAGAGAC	228		
QY	583	CATCAATCCTGCTAGAGTCAGGGTGGCGACGCCCAAGGGTGGCTCACCAGAGCTCGAG	642		
DB	227	CATCAATCCTGCTAGAGTCAGGGTGGCGACGCCCAAGGGTGGCTCACCAGAGCTCGAG	168		
QY	643	AGTCCTCTCATCTTCAGGTCCATTTACAGCTCTCTGGCATTTTAACTACACAGCATCCAGTGG	702		
DB	167	AGTCCTCTCATCTTCAGGTCCATTTACAGCTCTCTGGCATTTTAACTACACAGCATCCAGTGG	108		
QY	703	TCCCCAAAGGAATCCCTTTCTTAGCCTCTCTGACATGAGTCTGCTGGAAAGAGCATCCAAACA	762		

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 13:37:01 ; Search time 56 Seconds
(without alignments)
4413.955 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgggggaaatctgcc.....actcaatgcagacacaaaaa 806

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	2.5	12847	1	US-08-550-715-1
2	19	2.4	425	4	US-09-641-638-326
3	19	2.4	465	6	5496550-9
4	19	2.4	466	6	5496550-7
5	19	2.4	472	4	US-08-943-731-123
6	19	2.4	2249	2	US-08-463-081B-34
7	19	2.4	2249	2	US-08-461-379A-34
8	19	2.4	2249	2	US-08-462-390B-34
9	19	2.4	2249	3	US-08-463-074B-34
10	19	2.4	2249	3	US-08-465-585C-34
11	19	2.4	2249	3	US-08-652-446-34
12	19	2.4	2318	4	US-09-851-062-3
13	19	2.4	2362	1	US-08-285-087-1
14	19	2.4	2362	1	US-08-621-493-1
15	19	2.4	2362	2	US-08-965-688-1
16	19	2.4	2362	4	US-09-260-173-1
17	19	2.4	3865	1	US-08-832-883-48
18	19	2.4	3865	2	US-08-832-877-48
19	19	2.4	4281	4	US-09-357-206A-8
20	19	2.4	5630	2	US-08-937-931-1
21	19	2.4	5630	4	US-09-285-502-1
22	19	2.4	5630	4	US-09-709-126-1
23	19	2.4	5630	4	US-09-871-385A-1
24	19	2.4	7521	4	US-09-004-838-116
25	19	2.4	7874	4	US-09-780-175-96
26	19	2.4	9365	4	US-09-608-285A-8
27	19	2.4	9365	4	US-09-350-836B-8

c	28	19	2.4	9365	4	US-09-370-265-8	Sequence 8, Appli
	29	19	2.4	10881	4	US-09-357-206A-9	Sequence 9, Appli
	30	19	2.4	12286	4	US-09-357-206A-1	Sequence 1, Appli
c	31	19	2.4	14747	4	US-09-608-285A-42	Sequence 42, Appli
	32	19	2.4	14796	4	US-08-975-080-35	Sequence 35, Appli
	33	19	2.4	14796	4	US-09-630-706-10	Sequence 10, Appli
	34	19	2.4	14796	4	US-09-496-694B-3	Sequence 3, Appli
c	35	19	2.4	15977	4	US-09-608-285A-59	Sequence 59, Appli
	36	19	2.4	18073	4	US-09-078-294-12	Sequence 12, Appli
c	37	19	2.4	24183	4	US-08-943-731-3	Sequence 3, Appli
	38	19	2.4	72604	4	US-09-268-992-7	Sequence 7, Appli
	39	19	2.4	72604	4	US-09-557-474-7	Sequence 7, Appli
	40	19	2.4	98844	4	US-09-791-211-10	Sequence 10, Appli
c	41	19	2.4	162450	4	US-09-345-882-1	Sequence 1, Appli
c	42	18	2.2	21	4	US-09-165-868-8	Sequence 8, Appli
c	43	18	2.2	46	1	US-08-232-177A-349	Sequence 349, App
	44	18	2.2	66	1	US-08-232-177A-301	Sequence 301, App
	45	18	2.2	206	2	US-08-623-906A-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-08-550-715-1
; Sequence 1, Application US/08550715
; Patent No. 5750345
; GENERAL INFORMATION:
; APPLICANT: Bowie, Lemuel J.
; TITLE OF INVENTION: Human -Thalassemia Mutations as a Predictor of
; TITLE OF INVENTION: Blood-Related Disorders
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/550,715
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 28493/32834
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12847 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(6703..6797, 6915..7119, 7262..7387)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(10514..10608, 10726..10930, 11080..11205)
; US-08-550-715-1

Query Match 2.5%; Score 20; DB 1; Length 12847;
Best Local Similarity 100.0%; Pred. No. 1.8;

```

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 770 AATAAATAAATAAATAACT 789
    |||||
Db 4590 AATAAATAAATAAATAACT 4609

RESULT 2
US-09-641-638-326/c
; Sequence 326 Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilva
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641.638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 326
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 276
; OTHER INFORMATION: 10-65-276 : polymorphic base A or G
; NAME/KEY: misc_binding
; LOCATION: 256..275
; OTHER INFORMATION: 10-65-276.mis1, potential
; NAME/KEY: misc binding
; LOCATION: 277..295
; OTHER INFORMATION: 10-65-276.mis2, complement
; NAME/KEY: primer_bind
; LOCATION: 1..19
; OTHER INFORMATION: upstream amplification primer
; NAME/KEY: primer_bind
; LOCATION: 408..425
; OTHER INFORMATION: downstream amplification primer, complement
; NAME/KEY: misc_binding
; LOCATION: 264..288
; OTHER INFORMATION: 10-65-276 potential probe
US-09-641-638-326

Query Match 2.4%; Score 19; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 770 AATAAATAAATAAATAAC 788
    |||||
Db 350 AATAAATAAATAAATAAC 332

RESULT 3
5496550-9
; Patent No. 5496550
; APPLICANT: WALLACH, MICHAEL; PUGATSCH, THEA; MENCHER, DAVID
; TITLE OF INVENTION: METHOD OF REDUCING THE OUTPUT OF EIMERIA
; OOCYSTS FROM A NEWBORN CHICK
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/108,763
; FILING DATE: 17-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 642,219
; FILING DATE: 16-JAN-1991
; APPLICATION NUMBER: 310,603
; FILING DATE: 14-FEB-1989
; APPLICATION NUMBER: 155,245
; FILING DATE: 12-FEB-1988
; APPLICATION NUMBER: 896,611
; FILING DATE: 14-AUG-1986
; SEQ ID NO: 7
; LENGTH: 466
5496550-7

Query Match 2.4%; Score 19; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 769 TAATAAATAAATAAATAAA 787
    |||||
Db 204 TAATAAATAAATAAATAAA 222

RESULT 5
US-08-943-731-123/c
; Sequence 123, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES

```

```

; APPLICATION NUMBER: US/08/108,763
; FILING DATE: 17-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 642,219
; FILING DATE: 16-JAN-1991
; APPLICATION NUMBER: 310,603
; FILING DATE: 14-FEB-1989
; APPLICATION NUMBER: 155,245
; FILING DATE: 12-FEB-1988
; APPLICATION NUMBER: 896,611
; FILING DATE: 14-AUG-1986
; SEQ ID NO: 9
; LENGTH: 465
5496550-9

Query Match 2.4%; Score 19; DB 6; Length 465;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 769 TAATAAATAAATAAATAAA 787
    |||||
Db 203 TAATAAATAAATAAATAAA 221

RESULT 4
5496550-7
; Patent No. 5496550
; APPLICANT: WALLACH, MICHAEL; PUGATSCH, THEA; MENCHER, DAVID
; TITLE OF INVENTION: METHOD OF REDUCING THE OUTPUT OF EIMERIA
; OOCYSTS FROM A NEWBORN CHICK
; NUMBER OF SEQUENCES: 10
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/108,763
; FILING DATE: 17-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 642,219
; FILING DATE: 16-JAN-1991
; APPLICATION NUMBER: 310,603
; FILING DATE: 14-FEB-1989
; APPLICATION NUMBER: 155,245
; FILING DATE: 12-FEB-1988
; APPLICATION NUMBER: 896,611
; FILING DATE: 14-AUG-1986
; SEQ ID NO: 7
; LENGTH: 466
5496550-7

Query Match 2.4%; Score 19; DB 6; Length 466;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 769 TAATAAATAAATAAATAAA 787
    |||||
Db 204 TAATAAATAAATAAATAAA 222

RESULT 5
US-08-943-731-123/c
; Sequence 123, Application US/08943731
; Patent No. 6265157
; GENERAL INFORMATION:
; APPLICANT: PROCKOP, DARWIN J.
; APPLICANT: SPOTILA, LORETTA D.
; APPLICANT: DELTAS, CONSTANTINOS D.
; APPLICANT: SEREDA, LARISA
; APPLICANT: LARSON, ANDREA W.
; APPLICANT: PACK, MICHAEL
; APPLICANT: COLIGE, ALAIN
; APPLICANT: EARLY, JAMES
; APPLICANT: KORKKO, JARMO
; APPLICANT: ALA-KORKKO, LEENA, et al.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING
; TITLE OF INVENTION: ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES

```



```

; NUMBER OF SEQUENCES: 666
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITICH SCHWARZE JACOBS & NADEL, P.C.
; STREET: ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
; STREET: FLR.
; CITY: PHILADELPHIA
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,731
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/212,322
; FILING DATE: 14-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/803,628
; FILING DATE: 03-DEC-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: DOYLE LEARY PH.D., KATHRYN
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9598-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 123:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 472 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-943-731-123
;
; Query Match 2.4%; Score 19; DB 4; Length 472;
; Best Local Similarity 100.0%; Pred. No. 5.3;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy 769 TAATAAATAAATAAATAA 787
Db 106 TAATAAATAAATAAATAA 88
;
;
RESULT 6
US-08-463-081B-34
; Sequence 34, Application US/08463081B
; Patent No. 5871960
; Patent No. 5871960 5837487
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding C85 Polypeptide,
; TITLE OF INVENTION: Vector and Transformed Cell Thereof, and Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; STREET: 444 South Flower St. - Suite 1900
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: . Version #1.25

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,081B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,736
; FILING DATE: 10-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P66 38150 (DART-060)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2249 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-463-081B-34
;
; Query Match 2.4%; Score 19; DB 2; Length 2249;
; Best Local Similarity 100.0%; Pred. No. 5.4;
; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
Qy 770 AATAAATAAATAAATAAAC 788
Db 1817 AATAAATAAATAAATAAAC 1835
;
;
RESULT 7
US-08-461-379A-34
; Sequence 34, Application US/08461379A
; Patent No. 5871961
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding C85 Polypeptide,
; TITLE OF INVENTION: Vector and Transformed Cell Thereof, and
; TITLE OF INVENTION: Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; ADDRESSEE: (B) STREET: One Westlakes-Berwyn
; CITY: Valley Forge
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; SOFTWARE: Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,379A
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 08/330,108; 08/104,736
; APPLICATION NUMBER: & 07/796,066
; FILING DATE: 27-OCT-1994; 10-AUG-1993 & 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: DART-070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)470-0700
; TELEFAX: (610)470-0701
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:

```

```
; LENGTH: 2249 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-461-379A-34

Query Match 2.4%; Score 19; DB 2; Length 2249;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
Db 1817 AATAAATAAATAAATAAAC 1835

RESULT 8
US-08-462-390B-34
; Sequence 34, Application US/08462390B
; Patent No. 5882894
; GENERAL INFORMATION:
; APPLICANT: Smith, K. A., & Beadling, C.
; TITLE OF INVENTION: Nucleic Acids Encoding CR8 Polypeptide, Vector and
; TRANSFORMED CELL THEREOF, and Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; CITY: Valley Forge
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,390B
; FILING DATE: 5-JUNE-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/330,108
; FILING DATE: 27-OCT-1994
; APPLICATION NUMBER: USSN 08/104,736
; FILING DATE: 10-AUG-1993
; APPLICATION NUMBER: USSN 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: DART-040
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610)407-0700
; TELEFAX: (610)407-0701
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2249 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-462-390B-34

Query Match 2.4%; Score 19; DB 2; Length 2249;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
Db 1817 AATAAATAAATAAATAAAC 1835

RESULT 9
```

```
US-08-463-074B-34
; Sequence 34, Application US/08463074B
; Patent No. 6020155
; GENERAL INFORMATION:
; APPLICANT: Smith, Kendall A. & Beadling, Carol
; TITLE OF INVENTION: Nucleic Acids Encoding CRI Fusion Protein, Vector an
; TRANSFORMED CELL THEREOF, and Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0,
; Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,074B
; FILING DATE: 5-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/104,736
; FILING DATE: 10-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/796,066
; FILING DATE: 20-NOV-91
; ATTORNEY/AGENT INFORMATION:
; NAME: Viviana Amzel, Ph. D.
; REGISTRATION NUMBER: 30,930
; REFERENCE/DOCKET NUMBER: P66 38143 (DART-020)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 622-7700
; TELEFAX: (213) 489-4210
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2249 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-463-074B-34

Query Match 2.4%; Score 19; DB 3; Length 2249;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
Db 1817 AATAAATAAATAAATAAAC 1835

RESULT 10
US-08-465-585C-34
; Sequence 34, Application US/08465585C
; Patent No. 6027914
; GENERAL INFORMATION:
; APPLICANT: Smith, K. A., & Beadling, C.
; TITLE OF INVENTION: Nucleic Acids Encoding CR6 Polypeptide, Vecto
; TRANSFORMED CELL THEREOF, and Expression Thereof
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 900071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
US-08-465-585C-34
```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,585C
FILING DATE: 5-JUNE-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/330,108
FILING DATE: 27-OCT-1994
APPLICATION NUMBER: USSN 08/104,736
FILING DATE: 10-AUG-1993
APPLICATION NUMBER: USSN 07/796,066
FILING DATE: 20-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: P66 38149 (DART-050)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 4894210
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 2249 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-465-585C-34

Query Match 2.4%; Score 19; DB 3; Length 2249;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
|||||
Db 1817 AATAAATAAATAAATAAAC 1835

RESULT 11
US-08-652-446-34
Sequence 34, Application US/08652446
Patent No. 6057427
GENERAL INFORMATION:
APPLICANT: Smith, Kendall A. & Beadling, Carol
TITLE OF INVENTION: Nucleic Acids Encoding CRS
TITLE OF INVENTION: Polypeptide, Vector and Transformed Cell Thereof, and
TITLE OF INVENTION: Expression Thereof
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: PRETTY, SCHROEDER & POPLAWSKI
ADDRESS: (B) STREET:
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0,
SOFTWARE: Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/652,446
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP App. # 96921319.8
FILING DATE: 5-JAN-1998
APPLICATION NUMBER: PCT/US/96/09194
FILING DATE: 5-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/330,108
FILING DATE: 27-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,074

444 South Flower St. - Suite 1900

FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,337
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/462,390
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/465,585
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/463,081
FILING DATE: 5-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,379
FILING DATE: 5-JUN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/739,523
FILING DATE: 29-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Viviana Amzel, Ph. D.
REGISTRATION NUMBER: 30,930
REFERENCE/DOCKET NUMBER: EP66 40035
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 622-7700
TELEFAX: (213) 489-4210
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 2249 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-652-446-34

Query Match 2.4%; Score 19; DB 3; Length 2249;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
|||||
Db 1817 AATAAATAAATAAATAAAC 1835

RESULT 12
US-09-851-062-3/c
Sequence 3, Application US/09851062
Patent No. 6448081
GENERAL INFORMATION:
APPLICANT: Brenda F. Baker
TITLE OF INVENTION: ANTISENSE MODULATION OF INTERLEUKIN 12 P40 SUBUNIT EXPRESSION
FILE REFERENCE: RTS-0247
CURRENT APPLICATION NUMBER: US/09/851,062
CURRENT FILING DATE: 2001-05-07
NUMBER OF SEQ ID NOS: 87
SEQ ID NO 3
LENGTH: 2318
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (14)...(1000)
US-09-851-062-3

Query Match 2.4%; Score 19; DB 4; Length 2318;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
|||||
Db 2274 AATAAATAAATAAATAAAC 2256

RESULT 13

US-08-265-087-1/c
 ; Sequence 1, Application US/08265087
 ; Patent No. 5571515
 ; GENERAL INFORMATION:
 ; APPLICANT: Scott, Phillip
 ; APPLICANT: Trinchieri, Giorgio
 ; TITLE OF INVENTION: Compositions and Methods for Use of
 ; TITLE OF INVENTION: IL-12 as an Adjuvant
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Howson and Howson
 ; STREET: Spring House Corporate Center, PO Box 457
 ; CITY: Spring House
 ; STATE: Pennsylvania
 ; COUNTRY: USA
 ; ZIP: 19477
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/265,087
 ; FILING DATE:
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/229,282
 ; FILING DATE: 18-APR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bak, Mary E.
 ; REGISTRATION NUMBER: 31,215
 ; REFERENCE/DOCKET NUMBER: WST51AUSA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-540-9206
 ; TELEFAX: 215-540-5818
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2362 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 33..1016
 ; US-08-265-087-1

Query Match 2.4%; Score 19; DB 1; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
 |||||
 Db 2293 AATAAATAAATAAATAAAC 2275

RESULT 14

US-08-621-493-1/c
 ; Sequence 1, Application US/08621493
 ; Patent No. 5723127
 ; GENERAL INFORMATION:
 ; APPLICANT: Scott, Phillip
 ; APPLICANT: Trinchieri, Giorgio
 ; TITLE OF INVENTION: Compositions and Methods for Use of
 ; TITLE OF INVENTION: IL-12 as an Adjuvant
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Howson and Howson
 ; STREET: Spring House Corporate Center, PO Box 457
 ; CITY: Spring House
 ; STATE: Pennsylvania

; COUNTRY: USA
 ; ZIP: 19477
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/621,493
 ; FILING DATE: 25-MAR-1996
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/265,087
 ; FILING DATE: 17-JUN-1994
 ; APPLICATION NUMBER: US 08/229,282
 ; FILING DATE: 18-APR-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bak, Mary E.
 ; REGISTRATION NUMBER: 31,215
 ; REFERENCE/DOCKET NUMBER: WST51AUSA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-540-9206
 ; TELEFAX: 215-540-5818
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2362 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 33..1016
 ; US-08-621-493-1

Query Match 2.4%; Score 19; DB 1; Length 2362;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAATAAATAAAC 788
 |||||
 Db 2293 AATAAATAAATAAATAAAC 2275

RESULT 15

US-08-965-688-1/c
 ; Sequence 1, Application US/08965688
 ; Patent No. 5976539
 ; GENERAL INFORMATION:
 ; APPLICANT: Scott, Phillip
 ; APPLICANT: Trinchieri, Giorgio
 ; TITLE OF INVENTION: Compositions and Methods for Use of
 ; TITLE OF INVENTION: IL-12 as an Adjuvant
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Howson and Howson
 ; STREET: Spring House Corporate Center, PO Box 457
 ; CITY: Spring House
 ; STATE: Pennsylvania
 ; COUNTRY: USA
 ; ZIP: 19477
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/965,688
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/621,493
 ; FILING DATE: 25-MAR-1996

; APPLICATION NUMBER: 08/265,087
; FILING DATE: 17-JUN-1994
; APPLICATION NUMBER: US 08/229,282
; FILING DATE: 18-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST51AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2362 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 33..1016
; US-08-965-688-1

Query Match 2.4%; Score 19; DB 2; Length 2362;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 770 AATAAATAAATAAATAAAC 788
|||||
Db 2293 AATAAATAAATAAATAAAC 2275

Search completed: May 17, 2003, 17:13:39
Job time : 86 secs

THIS PAGE BLANK (user)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 14:37:12 ; Search time 116 Seconds
(without alignments)
8962.846 Million cell updates/sec

Title: US-09-724-000A-4
Perfect score: 806
Sequence: 1 ggaacgagggaaatctgcc.....actcaatgcagacacaaaa 806

Scoring table: OLIGO NUC

Gapop_60.0 , Gapext 60.0

Searched: 810007 seqs, 644969091 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1620014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	783	97.1	800	9	US-09-981-353-177
2	782	97.0	804	9	US-10-028-072-149
3	782	97.0	804	9	US-10-121-049-149
4	782	97.0	804	9	US-10-123-904-149
5	782	97.0	804	9	US-10-140-470-149
6	782	97.0	804	9	US-10-175-746-149
7	782	97.0	804	9	US-10-176-918-149
8	782	97.0	804	9	US-10-176-921-149
9	782	97.0	804	9	US-10-137-865-149
10	782	97.0	804	9	US-10-140-474-149
11	782	97.0	804	9	US-10-142-431-149
12	782	97.0	804	9	US-10-143-114-149
13	782	97.0	804	9	US-10-140-002-149
14	782	97.0	804	9	US-10-142-419-149
15	782	97.0	804	9	US-10-123-262-149
16	782	97.0	804	9	US-10-142-423-149
17	782	97.0	804	9	US-10-121-050-149
18	782	97.0	804	9	US-10-141-755-149
19	782	97.0	804	9	US-10-143-032-149

20	782	97.0	804	9	US-10-123-108-149	Sequence 149, App
21	782	97.0	804	9	US-10-123-236-149	Sequence 149, App
22	782	97.0	804	9	US-10-123-261-149	Sequence 149, App
23	782	97.0	804	9	US-10-140-921-149	Sequence 149, App
24	782	97.0	804	9	US-10-140-928-149	Sequence 149, App
25	782	97.0	804	9	US-10-121-045-149	Sequence 149, App
26	782	97.0	804	9	US-10-123-292-149	Sequence 149, App
27	782	97.0	804	9	US-10-123-903-149	Sequence 149, App
28	782	97.0	804	9	US-10-124-819-149	Sequence 149, App
29	782	97.0	804	9	US-10-124-822-149	Sequence 149, App
30	782	97.0	804	9	US-10-140-925-149	Sequence 149, App
31	782	97.0	804	9	US-10-160-498-149	Sequence 149, App
32	782	97.0	804	9	US-10-121-041-149	Sequence 149, App
33	782	97.0	804	9	US-10-121-043-149	Sequence 149, App
34	782	97.0	804	9	US-10-121-047-149	Sequence 149, App
35	782	97.0	804	9	US-10-123-215-149	Sequence 149, App
36	782	97.0	804	9	US-10-123-902-149	Sequence 149, App
37	782	97.0	804	9	US-10-123-908-149	Sequence 149, App
38	782	97.0	804	9	US-10-123-909-149	Sequence 149, App
39	782	97.0	804	9	US-10-123-910-149	Sequence 149, App
40	782	97.0	804	9	US-10-124-813-149	Sequence 149, App
41	782	97.0	804	9	US-10-124-817-149	Sequence 149, App
42	782	97.0	804	9	US-10-124-824-149	Sequence 149, App
43	782	97.0	804	9	US-10-125-922-149	Sequence 149, App
44	782	97.0	804	9	US-10-125-924-149	Sequence 149, App
45	782	97.0	804	9	US-10-127-825A-149	Sequence 149, App

ALIGNMENTS

RESULT 1
US-09-981-353-177
; Sequence 177, Application US/09981353
; Patent No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W.
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 177
; LENGTH: 800
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 1736965CB1
US-09-981-353-177

Query Match	97.1%	Score 783;	DB 9;	Length 800;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 783;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	10	GAATAATCGCTTCTCACCATGAGGCTTCTAGTCTTTCCAGCGCTCTCTGTATCTCTGCT	69	
Db	6	GAATAATCGCTTCTCACCATGAGGCTTCTAGTCTTTCCAGCGCTCTCTGTATCTCTGCT	65	
Qy	70	TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGAGCGCTCTGCCAAGGCTTGTCAGG	129	
Db	66	TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGAGCGCTCTGCCAAGGCTTGTCAGG	125	
Qy	130	CAGAGAACCGAGGCTCTGTGTCACCGAGTCCCTAGCCCAACTCAACAACTGAAAGG	189	
Db	126	CAGAGAACCGAGGCTCTGTGTCACCGAGTCCCTAGCCCAACTCAACAACTGAAAGG	185	
Qy	190	ACATCATGTAGGCTCTGTAACCATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	249	
Db	186	ACATCATGTAGGCTCTGTAACCATGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	245	

Qy 250 GCCTGGGCACTCCACAGGTGTAGCTCCCAAGCAAGACTCCAGACAGCGGAGAAC 309
Db 246 GCCTGGGCACTCCACAGGTGTAGCTCCCAAGCAAGACTCCAGACAGCGGAGAAC 305
Qy 310 TCATGCTGGCACTGAGGTACAGAGCCTCTGTCTCCCTTTAGCCTTACAGCA 369
Db 306 TCATGCTGGCACTGAGGTACAGAGCCTCTGTCTCCCTTTAGCCTTACAGCA 365
Qy 370 GTGAGCTCAATGTTGGAGGCTTCATCTCGGCTGCAAGGACCTGGGAAGTTCAGA 429
Db 366 GTGAGCTCAATGTTGGAGGCTTCATCTCGGCTGCAAGGACCTGGGAAGTTCAGA 425
Qy 430 ACTCACGCTCTGTCTCAATGTTGCCATCAACTTTTCAGAGCTATCATGAGCCAACTCA 489
Db 426 ACTCACGCTCTGTCTCAATGTTGCCATCAACTTTTCAGAGCTATCATGAGCCAACTCA 485
Qy 490 CCCACAGGCTCTAGTGCACCATGTTGGGCTCTCCAGTGCAAAACACCGAGCATTC 549
Db 486 CCCACAGGCTCTAGTGCACCATGTTGGGCTCTCCAGTGCAAAACACCGAGCATTC 545
Qy 550 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGTGCAGGTGG 609
Db 546 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGTGCAGGTGG 605
Qy 610 CAAGCACCAAGGCTGACCAAGACTGACAGAGTCTCTCCATCTTCAGGTCCATTCA 669
Db 606 CAAGCACCAAGGCTGACCAAGACTGACCAAGACTGACAGAGTCTCTCCATCTTCAGGTCCATTCA 665
Qy 670 GCCTCTGGCATTTAATCAGCATCAGATGCTGAGTCCCAAGAAATCTCTTCTAGCCTCC 729
Db 666 GCCTCTGGCATTTAATCAGCATCAGATGCTGAGTCCCAAGAAATCTCTTCTAGCCTCC 725
Qy 730 TCACATGAGTCTGCTGGAAGAGCATCAACAAACAGTAAATAAATAAATAACT 789
Db 726 TCACATGAGTCTGCTGGAAGAGCATCAACAAACAGTAAATAAATAAATAAATAACT 785
Qy 790 CAA 792
Db 786 CAA 788

RESULT 2

US-10-028-072-149
; Sequence 149, Application US/10028072

; Publication No. US20030004311A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tamas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang

TITLE OF INVENTION:

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/028.072

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/049911

; PRIOR FILING DATE: 1997-06-18

; PRIOR APPLICATION NUMBER: 60/056974

; PRIOR FILING DATE: 1997-08-26

; PRIOR APPLICATION NUMBER: 60/059113

; PRIOR FILING DATE: 1997-09-17

; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694

;; PRIOR FILING DATE: 1997-12-16
;; PRIOR APPLICATION NUMBER: 60/072320
;; PRIOR FILING DATE: 1998-01-23
;; PRIOR APPLICATION NUMBER: 60/073612
;; PRIOR FILING DATE: 1998-02-04
;; PRIOR APPLICATION NUMBER: 60/074086
;; PRIOR FILING DATE: 1998-02-09
;; PRIOR APPLICATION NUMBER: 60/074092
;; PRIOR FILING DATE: 1998-02-09
;; PRIOR APPLICATION NUMBER: 60/077791
;; PRIOR FILING DATE: 1998-03-12
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079663
;; PRIOR FILING DATE: 1998-02-27
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081695
;; PRIOR FILING DATE: 1998-04-14
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081818
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082999
;; PRIOR FILING DATE: 1998-04-24
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085149
;; PRIOR FILING DATE: 1998-05-12
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086414
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086430
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088730
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088741
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 19/98-06-11

;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090538
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07

Query Match 97.0%; Score 782; DB 9; Length 804;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTTTCAGGCTGCTGTATATCTTCTGCT 69
Db 23 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTTTCAGGCTGCTGTATATCTTCTGCT 82
Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGCGCTCTGCCAAGGCTTGGTCAGG 129
Db 83 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGCGCTCTGCCAAGGCTTGGTCAGG 142
Qy 130 CAGGAGAACCAAGGCTCTGCTGCCACCGAGTCCCTTAGCCCCCAACTCAACAACTTGAAGG 189
Db 143 CAGGAGAACCAAGGCTCTGCTGCCACCGAGTCCCTTAGCCCCCAACTCAACAACTTGAAGG 202
Qy 190 ACATCATGTGAGGCTCTGTAACCATGCAAGCTTGAGCCAGAGCCCGCTTGGGTGGT 249
Db 203 ACATCATGTGAGGCTCTGTAACCATGCAAGCTTGAGCCAGAGCCCGCTTGGGTGGT 262
Qy 250 GCCTGGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGAGAAC 309
Db 263 GCCTGGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGAGAAC 322
Qy 310 TCATGCTGGCACCTGAGGTACCCAGCAGGCTCTCTGCTCTCCCTTTTTCAGCCTTTCACAGCA 369
Db 323 TCATGCTGGCACCTGAGGTACCCAGCAGGCTCTCTGCTCTCCCTTTTTCAGCCTTTCACAGCA 382
Qy 370 GTGAGCTGCAATGTTGAGGGCTTTCATCTCGGGCTGCAAGAGCCCTGGGAAAGTTCCAGA 429
Db 383 GTGAGCTGCAATGTTGAGGGCTTTCATCTCGGGCTGCAAGAGCCCTGGGAAAGTTCCAGA 442
Qy 430 ACTCCAGCTCTGCTCTCAATTTGCCATCACTTTTCAGAGCTATCATGAGCCAACTCA 489
Db 443 ACTCCAGCTCTGCTCTCAATTTGCCATCACTTTTCAGAGCTATCATGAGCCAACTCA 502
Qy 490 CCCACAGGGCTCAGTGGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGCAGCATTTCC 549
Db 503 CCCACAGGGCTCAGTGGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGCAGCATTTCC 562
Qy 550 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTGTAGTGCAGGTGG 609
Db 563 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTGTAGTGCAGGTGG 622
Qy 610 CAAGCACCCAGGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCA 669
Db 623 CAAGCACCCAGGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCA 682

Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCTCACATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 69
Db 23 GAAATCTGCTTCTCACATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 82
Qy 70 TCTCTGCTTCTCCATCTTCTCACAGAGGAGGAGGCTCTGCAAGGCTGCTGAGG 129
Db 83 TCTCTGCTTCTCCATCTTCTCACAGAGGAGGAGGCTCTGCAAGGCTGCTGAGG 142
Qy 130 CAGGAGAACAGGCTCTGCTGCCACCGAGTCCCTAGTCTGCTTCCAGCTGCTGCTGAGG 189
Db 143 CAGGAGAACAGGCTCTGCTGCCACCGAGTCCCTAGTCTGCTTCCAGCTGCTGCTGAGG 202
Qy 190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTCCAGAGGAGGCTGCTGAGGCTGCTGAGG 249
Db 203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTCCAGAGGAGGCTGCTGAGGCTGCTGAGG 262
Qy 250 GCCTGGGCACTCCACAGGCTGAGTCTCCCAAGGAGGAGTCCAGAGGAGGAGGAGGAGG 309
Db 263 GCCTGGGCACTCCACAGGCTGAGTCTCCCAAGGAGGAGTCCAGAGGAGGAGGAGGAGG 322
Qy 310 TCATGCTGGCACTGAGTCTCCAGAGGCTGCTGCTGCTTCCAGCTGCTGCTGAGGAGGAGG 369
Db 323 TCATGCTGGCACTGAGTCTCCAGAGGCTGCTGCTGCTTCCAGCTGCTGCTGAGGAGGAGG 382
Qy 370 GTGAGCTGCAATGTTGGAGGCTTCTATCTGCGGCTGCAAGGAGGCTGCTGAGGAGGAGG 429
Db 383 GTGAGCTGCAATGTTGGAGGCTTCTATCTGCGGCTGCAAGGAGGCTGCTGAGGAGGAGG 442
Qy 430 ACTCCAGCTGCTGCTCAATGTTGCCATCAATCTTCCAGGAGTATCATGAGGAGGAGGAGG 489
Db 443 ACTCCAGCTGCTGCTCAATGTTGCCATCAATCTTCCAGGAGTATCATGAGGAGGAGGAGG 502
Qy 490 CCCACAGGCTGAGTCCGACCATGAGTGGGCTCTCCAGTGCAGGAGGAGGAGGAGGAGG 549
Db 503 CCCACAGGCTGAGTCCGACCATGAGTGGGCTCTCCAGTGCAGGAGGAGGAGGAGGAGG 562
Qy 550 ACCATGACCGGTGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 609
Db 563 ACCATGACCGGTGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 622
Qy 610 CAAGCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 669
Db 623 CAAGCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 682
Qy 670 GCCTCTGGCATTTAACTACAGGAGTCCAGTGGTCCCAAGGAGGAGGAGGAGGAGGAGG 729
Db 683 GCCTCTGGCATTTAACTACAGGAGTCCAGTGGTCCCAAGGAGGAGGAGGAGGAGGAGG 742

RESULT 5

US-10-140-470-149
; Sequence 149, Application US/10140470
; Publication No. US2003002231A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Geiritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-470-149

Query Match 97.0%; Score 782; DB 9; Length 804;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCTCACATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 69
Db 23 GAAATCTGCTTCTCACATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 82
Qy 70 TCTCTGCTTCTCCATCTTCTCACAGAGGAGGAGGCTCTGCAAGGCTGCTGAGG 129
Db 83 TCTCTGCTTCTCCATCTTCTCACAGAGGAGGAGGCTCTGCAAGGCTGCTGAGG 142
Qy 130 CAGGAGAACAGGCTCTGCTGCCACCGAGTCCCTAGTCTGCTTCCAGCTGCTCTGATCTGCT 189
Db 143 CAGGAGAACAGGCTCTGCTGCCACCGAGTCCCTAGTCTGCTTCCAGCTGCTCTGATCTGAT 202
Qy 190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTCCAGGAGGAGGAGGAGGAGGAGG 249
Db 203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTCCAGGAGGAGGAGGAGGAGGAGG 262
Qy 250 GCCTGGGCACTCCACAGGCTGAGTCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 309
Db 263 GCCTGGGCACTCCACAGGCTGAGTCTCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGG 322
Qy 310 TCATGCTGGCAGCTGAGGAGTACCCAGAGGAGGCTCTGCTGCTTCCAGCTTCCAGCTTCCAGCA 369
Db 323 TCATGCTGGCAGCTGAGGAGTACCCAGAGGAGGCTCTGCTGCTTCCAGCTTCCAGCTTCCAGCA 382
Qy 370 GTGAGCTGCAATGTTGGAGGCTTCTATCTGCGGCTGCAAGGAGGAGGAGGAGGAGGAGG 429
Db 383 GTGAGCTGCAATGTTGGAGGCTTCTATCTGCGGCTGCAAGGAGGAGGAGGAGGAGGAGG 442
Qy 430 ACTCCAGCTGCTGCTCAATGTTGCCATCAATCTTCCAGGAGTATCATGAGGAGGAGGAGG 489
Db 443 ACTCCAGCTGCTGCTCAATGTTGCCATCAATCTTCCAGGAGTATCATGAGGAGGAGGAGG 502
Qy 490 CCCACAGGCTGAGTCCGACCATGAGTGGGCTCTCCAGTGCAGGAGGAGGAGGAGGAGG 549
Db 503 CCCACAGGCTGAGTCCGACCATGAGTGGGCTCTCCAGTGCAGGAGGAGGAGGAGGAGG 562
Qy 550 ACCATGACCGGTGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 609
Db 563 ACCATGACCGGTGAGTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 622
Qy 610 CAAGCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 669
Db 623 CAAGCACCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 682
Qy 670 GCCTCTGGCATTTAACTACAGGAGTCCAGTGGTCCCAAGGAGGAGGAGGAGGAGGAGG 729
Db 683 GCCTCTGGCATTTAACTACAGGAGTCCAGTGGTCCCAAGGAGGAGGAGGAGGAGGAGG 742

Qy	730	TGACATGAGTCTGCTGGAAAAGAGCATCCAAACAACAAGTAATAAATAAATAACT	789
Db	743	TGACATGAGTCTGCTGGAAAAGAGCATCCAAACAACAAGTAATAAATAAATAACT	802
Qy	790	CA 791	
Db	803	CA 804	

RESULT 6

```

US-10-175-746-149
; Sequence 149, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-175-746-149

```

Db	383	GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCCCTGGGAAAGTTCGAGA	442
Qy	430	ACTCCAGTCTCTGTCTCAATTGTGCATCAACTTTTCAGAGCTATCATGAGCCAACTCTCA	489
Db	443	ACTCCAGTCTCTGTCTCAATTGTGCATCAACTTTTCAGAGCTATCATGAGCCAACTCTCA	502
Qy	490	CCCCACAGGGCCTCAGTCGCCACCATGTGGGCGCTCTCCAGTGCAAACCCAGCAGCATTC	549
Db	503	CCCCACAGGGCCTCAGTCGCCACCATGTGGGCGCTCTCCAGTGCAAACCCAGCAGCATTC	562
Qy	550	ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTGCTAGAGTGCAGGGTGG	609
Db	563	ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTGCTAGAGTGCAGGGTGG	622
Qy	610	CAAGCACCCGAAGGGTGGCTACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTGCCAATTC	669
Db	623	CAAGCACCCGAAGGGTGGCTACCAAGACTGCAGAGTCTCCTCCATCTTCAGGTGCCAATTC	682
Qy	670	GCCTCTCTGGCATTTAACTACCCAGCATTCAGTGGTCCCCAAGGAATCCCTTCTTAGCGCTCC	729
Db	683	GCCTCTCTGGCATTTAACTACCCAGCATTCAGTGGTCCCCAAGGAATCCCTTCTTAGCGCTCC	742
Qy	730	TGACATGAGTCTGCTGGAAAGAGCATCCAAAACAAACAAAGTAATAAATAAATAAATAA	789
Db	743	TGACATGAGTCTGCTGGAAAGAGCATCCAAAACAAACAAAGTAATAAATAAATAAATAA	802
Qy	790	CA 791	
Db	803	CA 804	

RESULT 7

US-10-176-918-149

; Sequence 149, Application US/10176918

; Publication No. US20030027275A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C382

; CURRENT APPLICATION NUMBER: US/10/176,918

; CURRENT FILING DATE: 2002-06-20

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 149

; LENGTH: 804

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-176-918-149

```

Query Match          97.0%; Score 782; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCCTTCTCACCATGAGGCTTCTAGTCCCTTTCACGCTGCTCTGTATCTGCT 69

```

Db	23	GAATAATCTGCGTCTCACCAATAGAGGCTTCTAGTCTCTTTTCAGCGCTGCTCTGTATCTCTGCT	82
Qy	70	TCTCTGCTTCTCCATCTTCTCCACAGAAGGAAGAGCGTCTCTGCCAAGGCCCTGGTCAGG	129
Db	83	TCTCTGCTTCTCCATCTTCTCCACAGAAGGAAGAGCGTCTCTGCCAAGGCCCTGGTCAGG	142
Qy	130	CAGGAGAACGAGGCTCTGCTGCCACACGAGTCCCTAGCGCCCACTCAACAAACCTGAAAGG	189
Db	143	CAGGAGAACGAGGCTCTGCTGCCACCGAGTCCCTAGCGCCCACTCAACAAACCTGAAAGG	202
Qy	190	ACATCATGTGAGGCTCTGTAAACCATGTCAAGCTTTGAGCCAGAGCCCGCGCTTTGGGTGGT	249
Db	203	ACATCATGTGAGGCTCTGTAAACCATGTCAAGCTTTGAGCCAGAGCCCGCGCTTTGGGTGGT	262
Qy	250	GCCTGGGGCACTCCACAGAGTGTAGCACTCCCAAAGCAAGACTCCACAGACGGGAGAACC	309
Db	263	GCCTGGGGCACTCCACAGAGTGTAGCACTCCCAAAGCAAGACTCCACAGACGGGAGAACC	322
Qy	310	TCATGCTGCGACCTGAGGTACCCAGCAGGCTCCTGTCTCTCCCTTTTCAGGCTTTCACAGCA	369
Db	323	TCATGCTGCGACCTGAGGTACCCAGCAGGCTCCTGTCTCTCCCTTTTCAGGCTTTCACAGCA	382
Qy	370	GTGAGCTGCAATGTTGGAGGGCTTTCATCTCGGGCTGCAAGGACCCCTGGGAAAGTTCCAGA	429
Db	383	GTGAGCTGCAATGTTGGAGGGCTTTCATCTCGGGCTGCAAGGACCCCTGGGAAAGTTCCAGA	442
Qy	430	ACTCCACGTCCTGTCTCAATGTGGCCATCAACTTTTCAGAGCTATCATGAGCCAACTTCA	489
Db	443	ACTCCACGTCCTGTCTCAATGTGGCCATCAACTTTTCAGAGCTATCATGAGCCAACTTCA	502
Qy	490	CCCCACAGGGGCTCAGTCGCCACCATGTGGGCGCTCTCCAGTGCMAACCAACCGAGCATTC	549
Db	503	CCCCACAGGGGCTCAGTCGCCACCATGTGGGCGCTCTCCAGTGCMAACCAACCGAGCATTC	562
Qy	550	ACCATGACCGGTACACAGTACAATAACAGAGACCATCAATCTCTGTCTAGAGTGCAGGGTGG	609
Db	563	ACCATGACCGGTACAGCTACAATAACAGAGACCATCAATCTCTGTCTAGAGTGCAGGGTGG	622
Qy	610	CAAGCACCCAAAGGGTGGCTGACCAAGATGACAGAGTCTCTCCATCTTCAGGTGCCATTCA	669
Db	623	CAAGCACCCAAAGGGTGGCTGACCAAGATGACAGAGTCTCTCCATCTTCAGGTGCCATTCA	682
Qy	670	GCCTCTCGGCATTTAACTACACAGCTCCAGTGGTCCCAAGGAATCCCTTCCTAGGCTCC	729
Db	683	GCCTCTCGGCATTTAACTACACAGCTCCAGTGGTCCCAAGGAATCCCTTCCTAGGCTCC	742
Qy	730	TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAAGTAATAAATAAATAAATAA	789
Db	743	TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAACAAAGTAATAAATAAATAAATAA	802
Qy	790	CA 791	
Db	803	CA 804	

RESULTS

US-100-176-921-149
; Sequence 149, Application US/10376921
; Publication No. US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Bresnini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.

Qy 790 CA 791
||
Db 803 CA 804

RESULT 9

US-10-137-865-149
; Sequence 149, Application US/10137865
; Publication No. US20030032155A1
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C154

; CURRENT APPLICATION NUMBER: US/10/137,865

; CURRENT FILING DATE: 2002-05-03

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 149

; LENGTH: 804

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-137-865-149

Query Match 97.0%; Score 782; DB 9; Length 804;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 69

Db 23 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 82

Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGAGGGCTCTGCGCAAGGCTGTGTCAGG 129

Db 83 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGAGGGCTCTGCGCAAGGCTGTGTCAGG 142

Qy 130 CAGGAGAACCGAGCTCTGCTGCCACCGAGTCCCTAGGCCCACTCAACAACTGAAAGG 189

Db 143 CAGGAGAACCGAGCTCTGCTGCCACCGAGTCCCTAGGCCCACTCAACAACTGAAAGG 202

Qy 190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGTGCT 249

Db 203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGTGCT 262

Qy 250 GCCTGGGGCACTCCACAGAGGTGTAGCACTCCCAAGCAAGAGCTCCAGACAGCGGAGAAC 309

Db 263 GCCTGGGGCACTCCACAGAGGTGTAGCACTCCCAAGCAAGAGCTCCAGACAGCGGAGAAC 322

Qy 310 TCATGCTGGCACTGAGTACCCAGAGCTCTGCTCTCCCTTTCCAGCTTCCAGCA 369

Db 323 TCATGCTGGCACTGAGTACCCAGAGCTCTGCTCTCCCTTTCCAGCTTCCAGCA 382

Qy 370 GTGAGCTCCAATGTTGGAGGCTTCATCTCGGCTGCAAGGACCTGCGGAAGTCCAGA 429

Db 383 GTGAGCTCCAATGTTGGAGGCTTCATCTCGGCTGCAAGGACCTGCGGAAGTCCAGA 442

Qy 430 ACTCCACGCTCTGTCTCAATTGTGCCATCAACTTTCCAGAGCTATCATGAGCCACCTCA 489

Db 443 ACTCCACGCTCTGTCTCAATTGTGCCATCAACTTTCCAGAGCTATCATGAGCCACCTCA 502

Qy 490 CCCACAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCAAAACCCGAGCATTC 549

Db 503 CCCACAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCAAAACCCGAGCATTC 562

Qy 550 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCCTCTAGAGTGCAAGGTTGG 609

Db 563 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCCTCTAGAGTGCAAGGTTGG 622

Qy 610 CAAGCACCCAGGGTGTGCTGACCAAGACTGCGAGAGTCTCTCCATCTTCAGGTCATTC 669

Db 623 CAAGCACCCAGGGTGTGCTGACCAAGACTGCGAGAGTCTCTCCATCTTCAGGTCATTC 682

Qy 670 GCCTCTGCGCATTTAACTACAGCATCCAGTGGTCCCAAGGATCCCTCTAGGCTCC 729

Db 683 GCCTCTGCGCATTTAACTACAGCATCCAGTGGTCCCAAGGATCCCTCTAGGCTCC 742

Qy 730 TGACATGAGTCTGCTGGAAGAGAGCATCCAAACAAACAGTAATAATAATAAACT 789

Db 743 TGACATGAGTCTGCTGGAAGAGAGCATCCAAACAAACAGTAATAATAATAAACT 802

Qy 790 CA 791

||

Db 803 CA 804

RESULT 10

US-10-140-474-149

; Sequence 149, Application US/10140474

; Publication No. US20030032156A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P330R1C162

; CURRENT APPLICATION NUMBER: US/10/140,474

; CURRENT FILING DATE: 2002-05-06

; Prior Application removed - See Palm or File Wrapper

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 149

; LENGTH: 804

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-140-474-149

Query Match 97.0%; Score 782; DB 9; Length 804;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 69

Db 23 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTTCCAGCTGCTCTGTATCTGCT 82

Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGAGGGCTCTGCGCAAGGCTGTGTCAGG 129

Db 83 TCTCTGCTTCTCCCATCTTCTCCACAGAGGGAAGAGGGCTCTGCCAAGGCTGGTCAGG 142
Qy 130 CAGGAGAACACAGGCTCTGTCGCCACCGAGTCCCTAGACCCCACTCAACAAACCTGAAAGG 189
Db 143 CAGGAGAACACAGGCTCTGTCGCCACCGAGTCCCTAGACCCCACTCAACAAACCTGAAAGG 202
Qy 190 ACATCATGTGAGGCTCTGTAACCATGCAAGCTTGAGCCAGAGCCCGCCCTTTGGGTGGT 249
Db 203 ACATCATGTGAGGCTCTGTAACCATGCAAGCTTGAGCCAGAGCCCGCCCTTTGGGTGGT 262
Qy 250 GCCTGGGGCACTCCACAGAGTGTAGCACTCCCAAGAGCAAGCTCCAGACAGCGGAGAAC 309
Db 263 GCCTGGGGCACTCCACAGAGTGTAGCACTCCCAAGAGCAAGCTCCAGACAGCGGAGAAC 322
Qy 310 TCATGCTGGCACTGAGGTACCCACAGAGCTCTGTCGCCCTTTCAGCCCTTCACAGCA 369
Db 323 TCATGCTGGCACTGAGGTACCCACAGAGCTCTGTCGCCCTTTCAGCCCTTCACAGCA 382
Qy 370 GTGAGCTGCAATGTGAGGCTTCATCTCGGGCTGCAAGACCTCGGAAAGTTCAGAG 429
Db 383 GTGAGCTGCAATGTGAGGCTTCATCTCGGGCTGCAAGACCTCGGAAAGTTCAGAG 442
Qy 430 ACTCCAGCTCTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCA 489
Db 443 ACTCCAGCTCTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCA 502
Qy 490 CCCACAGGGCTCAGTGGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGGAGCATTC 549
Db 503 CCCACAGGGCTCAGTGGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGGAGCATTC 562
Qy 550 ACCATGACCGGTACAGCTACAAATCCAGAGACCAATCAATCTGCTAGAGTGCAGGGTGG 609
Db 563 ACCATGACCGGTACAGCTACAAATCCAGAGACCAATCAATCTGCTAGAGTGCAGGGTGG 622
Qy 610 CAAGCACCAGGGGTGGTGCACCAAGCTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
Db 623 CAAGCACCAGGGGTGGTGCACCAAGCTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682
Qy 670 GCCTCTGGCATTTAACTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCCTCC 729
Db 683 GCCTCTGGCATTTAACTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCCTCC 742
Qy 730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAAGTAATAATAATAATAAACT 789
Db 743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAAGTAATAATAATAATAAACT 802
Qy 790 CA 791
Db 803 CA 804

RESULT 11
US-10-142-431-149
; Sequence 149, Application US/10142431
; Publication No. US20030036179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerriteen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P330R1C251
; CURRENT APPLICATION NUMBER: US/10/142,431
; CURRENT FILING DATE: 2002-05-10
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-142-431-149

Query Match 97.0%; Score 782; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTTCAGGCTGCTCTGTATCTCTGCT 69
Db 23 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTTCAGGCTGCTCTGTATCTCTGCT 82
Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGGAAGAGGCTCTGCAAGGCTGCTGTCAGG 129
Db 83 TCTCTGCTTCTCCATCTTCTCCACAGAGGGAAGAGGCTCTGCAAGGCTGCTGTCAGG 142
Qy 130 CAGGAGAACACAGGCTCTGCTGCCACCGAGTCCCTAGGCCCAACTCAACAAACCTGAAAGG 189
Db 143 CAGGAGAACACAGGCTCTGCTGCCACCGAGTCCCTAGGCCCAACTCAACAAACCTGAAAGG 202
Qy 190 ACATCATGTGAGGCTCTGTAACCATGCAAGCTTGAGCCAGAGCCCGCCCTTTGGGTGGT 249
Db 203 ACATCATGTGAGGCTCTGTAACCATGCAAGCTTGAGCCAGAGCCCGCCCTTTGGGTGGT 262
Qy 250 GCCTGGGGCACTCCACAGAGTGTAGCACTCCCAAGAGCAAGCTCCAGACAGCGGAGAAC 309
Db 263 GCCTGGGGCACTCCACAGAGTGTAGCACTCCCAAGAGCAAGCTCCAGACAGCGGAGAAC 322
Qy 310 TCATGCTGGCACTGAGGTACCCACAGAGCTCTGCTCTCCCTTTTTCAGCCCTTCACAGCA 369
Db 323 TCATGCTGGCACTGAGGTACCCACAGAGCTCTGCTCTCCCTTTTTCAGCCCTTCACAGCA 382
Qy 370 GTGAGCTGCAATGTGAGGCTTCATCTCGGGCTGCAAGACCTCGGAAAGTTCAGAG 429
Db 383 GTGAGCTGCAATGTGAGGCTTCATCTCGGGCTGCAAGACCTCGGAAAGTTCAGAG 442
Qy 430 ACTCCAGCTCTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCA 489
Db 443 ACTCCAGCTCTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCAATGTGCTCA 502
Qy 490 CCCACAGGGCTCAGTGGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGGAGCATTC 549
Db 503 CCCACAGGGCTCAGTGGCCACCATGTGGGCTCTCCAGTGCAAAACCCAGGAGCATTC 562
Qy 550 ACCATGACCGGTACAGCTACAAATCCAGAGACCAATCAATCTGCTAGAGTGCAGGGTGG 609
Db 563 ACCATGACCGGTACAGCTACAAATCCAGAGACCAATCAATCTGCTAGAGTGCAGGGTGG 622
Qy 610 CAAGCACCAGGGGTGGTGCACCAAGCTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 669
Db 623 CAAGCACCAGGGGTGGTGCACCAAGCTGCAGAGTCTCCTCCATCTTCAGGTCCATTCA 682
Qy 670 GCCTCTGGCATTTAACTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCCTCC 729
Db 683 GCCTCTGGCATTTAACTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCCTCC 742
Qy 730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAAGTAATAATAATAATAAACT 789
Db 743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAAGTAATAATAATAATAAACT 802
Qy 790 CA 791
Db 803 CA 804


```
RESULT 12
US-10-143-114-149
; Sequence 149, Application US/10143114
; Publication No. US20030036180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C211
; CURRENT APPLICATION NUMBER: US/10/143,114
; CURRENT FILING DATE: 2002-05-09
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-143-114-149

Query Match      97.0%; Score 782; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCCATCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 69
Db 23 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 82
Qy 70 TCTCTGCTTCTCCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 129
Db 83 TCTCTGCTTCTCCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 142
Qy 130 CAGGAGAACAGGCTTCTGCTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 189
Db 143 CAGGAGAACAGGCTTCTGCTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 202
Qy 190 ACATCATGTGAGGCTTCTGTAAACCATGCAAGCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 249
Db 203 ACATCATGTGAGGCTTCTGTAAACCATGCAAGCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 262
Qy 250 GCCTGGGGCACTCCACAGAGGCTTCTAGTCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 309
Db 263 GCCTGGGGCACTCCACAGAGGCTTCTAGTCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 322
Qy 310 TCATGCTGGGCACTCCACAGAGGCTTCTAGTCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 369
Db 323 TCATGCTGGGCACTCCACAGAGGCTTCTAGTCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 382
Qy 370 GTGAGCTGCAATGTGGAGGCTTCTATCTGGGCTGCAAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 429
Db 383 GTGAGCTGCAATGTGGAGGCTTCTATCTGGGCTGCAAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 442
Qy 430 ACTCCAGCTGCTTGTCTCAATGTGGCTTCTACATCTTCCAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 489
Db 443 ACTCCAGCTGCTTGTCTCAATGTGGCTTCTACATCTTCCAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 502
Qy 490 CCCCACAGGGGCTCAGTGGCCACCATGTGGGCTCTTCCAGTGTCAAAACCCAGGAGCATTTCC 549
```

```
Db 503 CCCACAGGGCTCAGTGGCCATCCAGTGGGCTCTCCAGTGCACACCCAGGATTTCC 562
Qy 550 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTAGTGTGAGGCTGG 609
Db 563 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTAGTGTGAGGCTGG 622
Qy 610 CAAGCACCCCAAGGGTGGTGCACCAAGACTGCAGAGTCTCTCCATCTTCCAGTCCATTCA 669
Db 623 CAAGCACCCCAAGGGTGGTGCACCAAGACTGCAGAGTCTCTCCATCTTCCAGTCCATTCA 682
Qy 670 GCCTCTGGCATTTAACTACACAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCTCC 729
Db 683 GCCTCTGGCATTTAACTACACAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCTCC 742
Qy 730 TGACATGAGTCTGCTGAAAGAGCATCCAAACAACAGTAATAAATAAATAAATAA 789
Db 743 TGACATGAGTCTGCTGAAAGAGCATCCAAACAACAGTAATAAATAAATAAATAA 802
Qy 790 CA 791
Db 803 CA 804

RESULT 13
US-10-140-002-149
; Sequence 149, Application US/10140002
; Publication No. US20030037623A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C59
; CURRENT APPLICATION NUMBER: US/10/140,002
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 149
; LENGTH: 804
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-002-149

Query Match      97.0%; Score 782; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 69
Db 23 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 82
Qy 70 TCTCTGCTTCTCCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 129
Db 83 TCTCTGCTTCTCCTCTTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 142
Qy 130 CAGGAGAACAGGCTTCTGCTCCACAGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTGCT 189
```


Db	143	CAGAGAAC	CAGGCTCTGCTG	CCACCAGTCCCTAG	CCCCAACTCAACAA	AACTGAAAG	202
Qy	190	ACATCATG	TGAGGCTCTGTA	AAACCATCAAGCTT	GTAGCCAGAGCCCG	CGCTTTGGTG	249
Db	203	ACATCATG	TGAGGCTCTGTA	AAACCATCAAGCTT	GAGCCAGAGCCCG	CGCTTTGGTG	262
Qy	250	GCCTGGGG	CACTCCACAGGT	GTAGCACTCCCA	AAAGCAAGACTT	CACAGCGGAGAAC	309
Db	263	GCCTGGGG	CACTCCACAGGT	GTAGCACTCCCA	AAAGCAAGACTT	CACAGCGGAGAAC	322
Qy	310	TCATGCTG	GCACCTGAGGT	ACCCAGCAGCCT	CTGTCTCCCTTT	CAGCTTTCACAGCA	369
Db	323	TCATGCTG	GCACCTGAGGT	ACCCAGCAGCCT	CTGTCTCCCTTT	CAGCTTTCACAGCA	382
Qy	370	GTGAGCTG	CAATGTGTGAG	GGGCTTCATCT	CGGGTGCAAGG	ACCTTGGGAAAGTTC	429
Db	383	GTGAGCTG	CAATGTGTGAG	GGGCTTCATCT	CGGGTGCAAGG	ACCTTGGGAAAGTTC	442
Qy	430	ACTCCAGT	CTCTGTCTCA	ATTGTGCCATCA	ACTTTTCAGAGCT	TATCATCAGGCCAAC	489
Db	443	ACTCCAGT	CTCTGTCTCA	ATTGTGCCATCA	ACTTTTCAGAGCT	TATCATCAGGCCAAC	502
Qy	490	CCCCACAG	GGCCTCAGT	TCGCCACCATG	TGGGCGCTCTCC	AGTGCAAACACCCAGCAGCATTC	549
Db	503	CCCCACAG	GGCCTCAGT	TCGCCACCATG	TGGGCGCTCTCC	AGTGCAAACACCCAGCAGCATTC	562
Qy	550	ACCATGAC	CCGGTCACAGT	TACAAATCC	CAGAGACCATCA	TCTTGCTAGAGTGCAGGGTGG	609
Db	563	ACCATGAC	CCGGTCACAGT	TACAAATCC	CAGAGACCATCA	TCTTGCTAGAGTGCAGGGTGG	622
Qy	610	CAAGCAC	CCCCAAGGGTGG	CTGACCAAGACT	TCAGAGTCTCT	CTCCATCTTCAGGTGCATTC	669
Db	623	CAAGCAC	CCCCAAGGGTGG	CTGACCAAGACT	TCAGAGTCTCT	CTCCATCTTCAGGTGCATTC	682
Qy	670	GCCTCTG	GGAATTTAACT	ACACGATC	CAGTGTGCC	CCAGGAATCCCTTCTAGCTCC	729
Db	683	GCCTCTG	GGAATTTAACT	ACACGATC	CAGTGTGCC	CCAGGAATCCCTTCTAGCTCC	742
Qy	730	TGACATCAG	TCTGCTG	AAAGAGCATC	CAAAACAA	CAAGTAATAATAATAATAA	789
Db	743	TGACATCAG	TCTGCTG	AAAGAGCATC	CAAAACAA	CAAGTAATAATAATAATAA	802
Qy	790	CA	791				
Db	803	CA	804				

RESULT 14

US-10-142-419-149
; Application US/10142419
; Publication No. US20030044945A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P33301C244
; CURRENT APPLICATION NUMBER: US/10/142.419

; Publication No. US20030049816A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Beresini, Maureen

; APPLICANT: DeForge, Laura

; APPLICANT: Desnoyers, Luc

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Sherwood, Steven

; APPLICANT: Smith, Victoria

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Watanabe, Colin K

; APPLICANT: Wood, William

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3330R1C38

; CURRENT APPLICATION NUMBER: US/10/123.262

; CURRENT FILING DATE: 2002-04-15

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 550

; SEQ ID NO 149

; LENGTH: 804

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-123-262-149

Query Match 97.0%; Score 782; DB 9; Length 804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 GAAATCTGCTTCTCACCATGAGGCTTTCTAGTCTCTTCCAGCCTGCTCTGTATCTCTGCT 69

Db 23 GAAATCTGCTTCTCACCATGAGGCTTTCTAGTCTCTTCCAGCCTGCTCTGTATCTCTGCT 82

Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGGAAGAGGCTCTGCGCAAGGCTGTGTCAGG 129

Db 83 TCTCTGCTTCTCCATCTTCTCCACAGAGGGAAGAGGCTCTGCGCAAGGCTGTGTCAGG 142

Qy 130 CAGGAGAACCGGCTCTGCTCCACCGAGTCCCTAGCCCACTCAACAAACCTGAAGG 189

Db 143 CAGGAGAACCGGCTCTGCTCCACCGAGTCCCTAGCCCACTCAACAAACCTGAAGG 202

Qy 190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGT 249

Db 203 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGT 262

Qy 250 GCCTGGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAAC 309

Db 263 GCCTGGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAAC 322

Qy 310 TCATGCTGGCACTCCAGAGTACCCAGAGCCTCTGCTCTCCCTTTAGCCCTTCACAGCA 369

Db 323 TCATGCTGGCACTCCAGAGTACCCAGAGCCTCTGCTCTCCCTTTAGCCCTTCACAGCA 382

Qy 370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGA 429

Db 383 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGA 442

Qy 430 ACTCCAGCTCTGCTCTCAATGTTGCCATCAACTTTTCAGAGCTATCATGAGCCCAACCTCA 489

Db 443 ACTCCAGCTCTGCTCTCAATGTTGCCATCAACTTTTCAGAGCTATCATGAGCCCAACCTCA 502

Qy 490 CCCACAGGGGCTCAGTCCGACCATGTGGGCTCTCCAGTGCACCAACCCAGGCAATCC 549

Db 503 CCCACAGGGGCTCAGTCCGACCATGTGGGCTCTCCAGTGCACCAACCCAGGCAATCC 562

Qy 550 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTGTAGAGTGCAGGGTGG 609

Db 563 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCTCTGTAGAGTGCAGGGTGG 622

Qy 610 CAAGCACCCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCA 669

Db 623 CAAGCACCCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCA 682

Qy 670 GCCTCTCTGGCATTTAACTTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCTCC 729

Db 683 GCCTCTCTGGCATTTAACTTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCTCC 742

Qy 730 TGACATGAGTCTGCTGGAAGAGCATCCAAACAAAGTAATAATAATAATAATAAACT 789

Db 743 TGACATGAGTCTGCTGGAAGAGCATCCAAACAAAGTAATAATAATAATAATAAACT 802

Qy 790 CA 791

Db 803 CA 804

Search completed: May 17, 2003, 17:20:47

Job time : 121 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 17, 2003, 11:55:46 ; Search time 2635 Seconds
(without alignments)
8902.035 Million cell updates/sec

Title: US-09-724-000A-4

Perfect score: 806

Sequence: 1 ggaacgagggaataatgtcc.....actcaatgcagacacaaaaa 806

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_on:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_on:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rod:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	806	100.0	806	6	AX342218	Sequence
2	782	97.0	804	6	AX464016	Sequence
3	616	76.4	198481	9	AC022389	Homo sapi
4	537	66.6	801	6	AX027767	Sequence
5	498	61.8	797	6	AX027773	Sequence
6	480	59.6	2063	9	AX025416	Homo sapi
7	323	40.1	485	6	AX351259	Sequence
8	299	37.1	382	6	AX330090	Sequence
9	196	24.3	366	6	AX335580	Sequence
10	196	24.3	366	6	AX335580	Sequence
11	196	24.3	366	6	AX408291	Sequence
12	93	11.5	4159	6	AX342222	Sequence
13	27	3.3	180176	10	AL671190	Mouse DNA
14	25	3.1	93165	9	AL357974	Human DNA
15	25	3.1	152245	2	AC022829	Homo sapi
16	25	3.1	153311	2	AC107297	Homo sapi
17	25	3.1	172043	9	AC078828	Homo sapi
18	25	3.1	193903	2	AC091156	Homo sapi
19	25	3.1	198057	2	AC073719	Mus muscu
20	25	3.1	231443	2	AC073691	Mus muscu
21	25	3.1	246962	2	AC073811	Mus muscu
22	24	3.0	94737	2	AC126011	Medicago
23	24	3.0	144818	9	AL139343	Human DNA
24	24	3.0	157758	9	AC098824	Homo sapi
25	24	3.0	162486	2	AC044779	Homo sapi
26	24	3.0	165114	2	AC098323	Rattus no
27	24	3.0	174194	2	AL845359	Danio rer
28	24	3.0	180718	2	AL844538	Mus muscu
29	24	3.0	192519	10	MM0278435	
30	24	3.0	202095	2	AC130044	Rattus no
31	24	3.0	215516	2	AC024694	Mus muscu
32	24	3.0	221647	10	AL591003	Mouse DNA
33	24	3.0	224446	2	AC122109	Mus muscu
34	23	2.9	7355	6	AX344889	Sequence
35	23	2.9	57276	2	AC124271	Homo sapi
36	23	2.9	57276	2	AC124271	Homo sapi
37	23	2.9	59438	2	AC090314	Homo sapi
38	23	2.9	64886	2	AC101145	Mus muscu
39	23	2.9	64886	2	AC101145	Mus muscu
40	23	2.9	66969	9	AL161789	Human DNA
41	23	2.9	72259	2	AC100677	Mus muscu
42	23	2.9	76396	2	AC012995	Drosophil
43	23	2.9	80922	2	AC097830	Rattus no
44	23	2.9	86945	2	AC002490	Homo sapi
45	23	2.9	88806	9	AL603890	Human DNA

ALIGNMENTS

RESULT 1	AX342218	Sequence 4 from Patent WO0198497.	806 bp	DNA	linear	PAT 12-JAN-2002
LOCUS	AX342218	Sequence 4 from Patent WO0198497.				
DEFINITION	AX342218	Sequence 4 from Patent WO0198497.				
ACCESSION	AX342218	Sequence 4 from Patent WO0198497.				
VERSION	AX342218.1	GI:18151763				
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
REFERENCE	1					
AUTHORS	Polverino, A.J. and Luethy, R.					
TITLE	Secreted epithelial colon stromal-1 polypeptides, nucleic acids encoding the same and uses thereof					


```
Db 563 ACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGGTGG 622
Qy 610 CAAGCACCACCAAGGTGGTGCAGAGACTGCAGAGTCTCTCTCCATCTTTCAGGTCCATTCA 669
Db 623 CAAGCACCACCAAGGTGGTGCAGAGACTGCAGAGTCTCTCTCCATCTTTCAGGTCCATTCA 682
Qy 670 GCCTCTGGCATTTAACTACAGAGATCCAGTGGTCCCAAGGAATCCCTTCTAGCTCC 729
Db 683 GCCTCTGGCATTTAACTACAGAGATCCAGTGGTCCCAAGGAATCCCTTCTAGCTCC 742
Qy 730 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAAGAGTAATAATAATAATAACT 789
Db 743 TGACATGAGTCTGCTGGAAAGAGCATCCAAACAAAGAGTAATAATAATAATAACT 802
Qy 790 CA 791
Db 803 CA 804

RESULT 3
AC022389
LOCUS Homo sapiens chromosome 10 clone RP11-124L5, complete sequence.
DEFINITION AC022389
ACCESSION AC022389.9 GI:21321764
VERSION HTG.
KEYWORDS SOURCE
ORGANISM human.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Smith, D.R.
JOURNAL Genome Therapeutics Corporation Sequencing Center: Human Genome
Unpublished
REFERENCE 2 (bases 1 to 198481)
AUTHORS Smith, D.R.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 3 (bases 1 to 198481)
AUTHORS Smith, D.R.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 4 (bases 1 to 198481)
AUTHORS Smith, D.R.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 5 (bases 1 to 198481)
AUTHORS Smith, D.R.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 6 (bases 1 to 198481)
AUTHORS Smith, D.R.
TITLE Direct Submission
JOURNAL Submitted (21-JUN-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
REFERENCE 7 (bases 1 to 198481)
AUTHORS Smith, D.R.
TITLE Direct Submission
JOURNAL Submitted (04-JUL-2002) Genome Therapeutics Corporation, 100 Beaver
Street, Waltham, MA 02453, USA
COMMENT On Jun 4, 2002 this sequence version replaced gi:19071578.
FEATURES
source Location/Qualifiers
1..198481
/db_xref="taxon:9606"
/organism="Homo sapiens"
/chromosome="10"
/clone="RP11-124L5"
```

```
BASE COUNT 55275 a 44232 c 45110 g 53864 t
ORIGIN
Query Match 76.4%; Score 616; DB 9; Length 198481;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 616; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 187 AGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTTGAGCCAGAGCCCGCCTTTGGGT 246
Db 43093 AGGACATCATGTGAGGCTCTGTAAACCATGCAAGCTTTGAGCCAGAGCCCGCCTTTGGGT 43152
Qy 247 GGTGCTCTGGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACACCGGAGA 306
Db 43153 GGTGCTCTGGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACACCGGAGA 43212
Qy 307 ACCTCATGCTGGCACCCTGAGGTACCCAGCAGCCTCTGCTCTCCCTTTGAGCCTTCACA 366
Db 43213 ACCTCATGCTGGCACCCTGAGGTACCCAGCAGCCTCTGCTCTCCCTTTGAGCCTTCACA 43272
Qy 367 GCAGTGAGCTGCAATGTTGGAGGCTTTCATCTCGGGCTGCAAGGACCTCTGGGAAAGTTCC 426
Db 43273 GCAGTGAGCTGCAATGTTGGAGGCTTTCATCTCGGGCTGCAAGGACCTCTGGGAAAGTTCC 43332
Qy 427 AGAATCCAGCTCTCTGCTCAATTGTGCCATCAACTTTTCAGAGCTATCATGAGCCCAACC 486
Db 43333 AGAATCCAGCTCTCTGCTCAATTGTGCCATCAACTTTTCAGAGCTATCATGAGCCCAACC 43392
Qy 487 TCACCCACAGGGCTCAGTCGCCACCATGTGGSCCTCTCCAGTGCAAAACACCCAGCAT 546
Db 43393 TCACCCACAGGGCTCAGTCGCCACCATGTGGSCCTCTCCAGTGCAAAACACCCAGCAT 43452
Qy 547 TCACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGG 606
Db 43453 TCACCATGACCGGTACAGCTACAAATCCAGAGACCATCAATCCTGCTAGAGTGCAGGG 43512
Qy 607 TGGCAAGCACCCAGGGTGGCTGACCAAGACTGTCAGAGTCTCTCCATCTTTCAGGTCCAT 666
Db 43513 TGGCAAGCACCCAGGGTGGCTGACCAAGACTGTCAGAGTCTCTCCATCTTTCAGGTCCAT 43572
Qy 667 TCAGCTCTCTGGCATTTAACTTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCC 726
Db 43573 TCAGCTCTCTGGCATTTAACTTACCAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCC 43632
Qy 727 TCCTGACATGAGTCTGCTGAAAGAGCATCCAAACAAAGTAATAATAATAATAATAA 786
Db 43633 TCCTGACATGAGTCTGCTGAAAGAGCATCCAAACAAAGTAATAATAATAATAATAA 43692
Qy 787 ACTCAATGCAGACACA 802
Db 43693 ACTCAATGCAGACACA 43708

RESULT 4
AX027767/c
LOCUS AX027767 801 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 1 from Patent WO0043509.
ACCESSION AX027767
VERSION AX027767.1 GI:10188619
KEYWORDS human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 801)
AUTHORS Vinals-Bassols, C.
TITLE Novel compounds
JOURNAL Patent: WO 0043509-A 1 27-JUL-2000;
SMITHKLINE BEECHAM BIOLOG (BE) ; VINALS BASSOLS CARLOTA (BE)
FEATURES
source Location/Qualifiers
1..801
/db_xref="taxon:9606"
/organism="Homo sapiens"
/clone="9606"
```

BASE COUNT 165 a 181 c 255 g 200 t
ORIGIN

Query Match 66.6%; Score 537; DB 6; Length 801;
Best Local Similarity 100.0%; Pred. No. 1.3e-312;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 261 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAACTCATGCTGGC 320
Db 537 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAACTCATGCTGGC 478

Qy 321 ACCTGAGGTACCCAGACGCTCTGTCTCCCTTTTACAGCTTTCACAGAGTGCAG 380
Db 477 ACCTGAGGTACCCAGACGCTCTGTCTCCCTTTTACAGCTTTCACAGAGTGCAG 418

Qy 381 TGTGGAGGGCTTCACTCTGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGTCC 440
Db 417 TGTGGAGGGCTTCACTCTGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGTCC 358

Qy 441 TTGTCTCAATTGTGCATCACTTTTACAGCTATCATGAGCCAACTCCACCCACAGGCG 500
Db 357 TTGTCTCAATTGTGCATCACTTTTACAGCTATCATGAGCCAACTCCACCCACAGGCG 298

Qy 501 CTCAGTCCGCCACCATGTGGGCTCTCCAGTGCAAACCCAGGACATTCACCATGACCGG 560
Db 297 CTCAGTCCGCCACCATGTGGGCTCTCCAGTGCAAACCCAGGACATTCACCATGACCGG 238

Qy 561 TCACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGTGCAGGTTCCAGCTTCTGGCA 620
Db 237 TCACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGTGCAGGTTCCAGCTTCTGGCA 178

Qy 621 GGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGCTTCTGGCA 680
Db 177 GGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGCTTCTGGCA 118

Qy 681 TTTAACTACCAAGCTCCAGTGGTCCCCAAGGAATCCCTTCTAGCTTCTGACATGATC 740
Db 117 TTTAACTACCAAGCTCCAGTGGTCCCCAAGGAATCCCTTCTAGCTTCTGACATGATC 58

Qy 741 TGCTGGAAGAGCATCCAAACCAAGTAATAATAATAATAATAATAATAATAATAATGAG 797
Db 57 TGCTGGAAGAGCATCCAAACCAAGTAATAATAATAATAATAATAATAATAATAATGAG 1

RESULT 5
AX027773/c
LOCUS AX027773 797 bp DNA linear PAT 16-SEP-2000
DEFINITION Sequence 7 from Patent WO0043509.
ACCESSION AX027773
VERSION AX027773.1 GI:10188625
KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 797)
AUTHORS Vinals-Bassols, C.
TITLE Novel compounds
JOURNAL Patent: WO 0043509-A 7 27-JUL-2000;
SMITHKLINE BEECHAM BIOLOG (BE) ; VINALS BASSOLS CARLOTA (BE)
FEATURES
Location/Qualifiers
source 1..797
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 164 a 180 c 254 g 199 t
ORIGIN

Query Match 61.8%; Score 498; DB 6; Length 797;
Best Local Similarity 100.0%; Pred. No. 5.1e-289;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 261 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAACTCATGCTGGC 320
Db 498 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAACTCATGCTGGC 320

Db 533 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAACTCATGCTGGC 474

Qy 321 ACCTGAGGTACCCAGACGCTCTGTCTCCCTTTTACAGCTTTCACAGAGTGCAGTCAA 380
Db 473 ACCTGAGGTACCCAGACGCTCTGTCTCCCTTTTACAGCTTTCACAGAGTGCAGTCAA 414

Qy 381 TGTGGAGGGCTTCACTCTGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGTCC 440
Db 413 TGTGGAGGGCTTCACTCTGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCACGTCC 354

Qy 441 TTGTCTCAATTGTGCCATCACTTTTACAGCTATCATGAGCCAACTCCACCCACAGGCG 500
Db 353 TTGTCTCAATTGTGCCATCACTTTTACAGCTATCATGAGCCAACTCCACCCACAGGCG 294

Qy 501 CTCAGTCCGCCACCATGTGGGCTCTCCAGTGCAAACCCAGGACATTCACCATGACCGG 560
Db 293 CTCAGTCCGCCACCATGTGGGCTCTCCAGTGCAAACCCAGGACATTCACCATGACCGG 234

Qy 561 TCACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGTGCAGGTTCCAGCTTCTGGCA 620
Db 233 TCACAGCTACAAATCCAGAGACCATCAATCTGCTAGAGTGCAGGTTCCAGCTTCTGGCA 174

Qy 621 GGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTTCAGGTCCATTCAGCTTCTGGCA 680
Db 173 GGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTTCAGGTCCATTCAGCTTCTGGCA 114

Qy 681 TTTAACTACCAAGCTCCAGTGGTCCCCAAGGAATCCCTTCTAGCTTCTGACATGATC 740
Db 113 TTTAACTACCAAGCTCCAGTGGTCCCCAAGGAATCCCTTCTAGCTTCTGACATGATC 54

Qy 741 TGCTGGAAGAGCATCCA 758
Db 53 TGCTGGAAGAGCATCCA 36

RESULT 6
AK025416 2063 bp mRNA linear PRI 29-SEP-2000
LOCUS AK025416
DEFINITION Homo sapiens cDNA: FLJ21763 fis, clone COLF6967.
ACCESSION AK025416
VERSION AK025416.1 GI:10437924
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens colon mucosa cDNA to mRNA, clone lib:ColF clone:COLF6967.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)
AUTHORS Watanabe, K., Kumagai, A., Itakura, S., Yamazaki, M., Tashiro, H., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2063)
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T., Isogai, T., Shibahara, T., Tanaka, T. and Nakamura, Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-3, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing; Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES
Location/Qualifiers
source 1..2063
/organism="Homo sapiens"
/db_xref="taxon:9606"


```
RESULT 9
AX330090/c
LOCUS AX330090 366 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 599 from Patent WO0194629.
ACCESSION AX330090
VERSION AX330090.1 GI:18103068
KEYWORDS human.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 599 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
source
1..366
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 77 a 80 c 119 g 90 t
ORIGIN
Query Match 24.3%; Score 196; DB 6; Length 366;
Best Local Similarity 99.6%; Pred. No. 2.3e-106;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 357 AGCCTTCACAGCAGTGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 416
Db 366 AGCCTTCACAGCAGTGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 307

Qy 417 GGAAGTTCAGAACTCCAGTCCCTTCTCAATTGTCATCAATTTGAGAGCTATCA 476
Db 306 GGAAGTTCAGAACTCCAGTCCCTTCTCAATTGTCATCAATTTGAGAGCTATCA 247

Qy 477 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAAC 536
Db 246 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAAC 187

Qy 537 CACCGAGCATTCACCATGACCGGTCAAGTCAAGTCCAGAGACCATCAATCCTGCTA 596
Db 186 CACCGAGCATTCACCATGACCGGTCAAGTCAAGTCCAGAGACCATCAATCCTGCTA 127

Qy 597 GAGTGCA 603
Db 126 GAGTGCA 120

RESULT 10
AX335580/c
LOCUS AX335580 366 bp DNA linear PAT 09-JAN-2002
DEFINITION Sequence 6089 from Patent WO0194629.
ACCESSION AX335580
VERSION AX335580.1 GI:18126299
KEYWORDS human.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Young, P.E., Augustus, M., Carter, K.C., Ebner, R., Endress, G.,
Horrigan, S., Soppet, D.R. and Weaver, Z.
TITLE Cancer gene determination and therapeutic screening using signature
gene sets
JOURNAL Patent: WO 0194629-A 6089 13-DEC-2001;
Avalon Pharmaceuticals (US)
FEATURES
Location/Qualifiers
source
1..366
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 77 a 80 c 119 g 90 t
ORIGIN
Query Match 24.3%; Score 196; DB 6; Length 366;
Best Local Similarity 99.6%; Pred. No. 2.3e-106;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 357 AGCCTTCACAGCAGTGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 416
Db 366 AGCCTTCACAGCAGTGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 307

Qy 417 GGAAGTTCAGAACTCCAGTCCCTTCTCAATTGTCATCAATTTGAGAGCTATCA 476
Db 306 GGAAGTTCAGAACTCCAGTCCCTTCTCAATTGTCATCAATTTGAGAGCTATCA 247

Qy 477 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAAC 536
Db 246 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAAC 187

Qy 537 CACCGAGCATTCACCATGACCGGTCAAGTCAAGTCCAGAGACCATCAATCCTGCTA 596
Db 186 CACCGAGCATTCACCATGACCGGTCAAGTCAAGTCCAGAGACCATCAATCCTGCTA 127

Qy 597 GAGTGCA 603
Db 126 GAGTGCA 120

RESULT 11
AX408291/c
LOCUS AX408291 366 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 938 from Patent WO0229103.
ACCESSION AX408291
VERSION AX408291.1 GI:21440996
KEYWORDS human.
SOURCE Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Alvares, C., Horne, D., Peres-da-Silva, S. and Vockley, J.G.
TITLE Gene expression profiles in liver cancer
JOURNAL Patent: WO 0229103-A 938 11-APR-2002;
GENE LOGIC INC (US)
FEATURES
Location/Qualifiers
source
1..366
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 77 a 80 c 119 g 90 t
ORIGIN
Query Match 24.3%; Score 196; DB 6; Length 366;
Best Local Similarity 99.6%; Pred. No. 2.3e-106;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 357 AGCCTTCACAGCAGTGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 416
Db 366 AGCCTTCACAGCAGTGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 307

Qy 417 GGAAGTTCAGAACTCCAGTCCCTTCTCAATTGTCATCAATTTGAGAGCTATCA 476
Db 306 GGAAGTTCAGAACTCCAGTCCCTTCTCAATTGTCATCAATTTGAGAGCTATCA 247

Qy 477 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAAC 536
Db 246 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAAC 187

Qy 537 CACCGAGCATTCACCATGACCGGTCAAGTCAAGTCCAGAGACCATCAATCCTGCTA 596
Db 186 CACCGAGCATTCACCATGACCGGTCAAGTCAAGTCCAGAGACCATCAATCCTGCTA 127

Qy 597 GAGTGCA 603
Db 126 GAGTGCA 120
```


Db 126 GAGTCA 120

RESULT 12
AX342222
LOCUS
DEFINITION Sequence 8 from Patent WO0198497.
ACCESSION AX342222
VERSION AX342222.1 GI:18151765
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Polverino, A.J. and Luethy, R.
TITLE Secreted epithelial colon stromal-1 polypeptides, nucleic acids
encoding the same and uses thereof
JOURNAL Patent: WO 0198497-A 8 27-DEC-2001;
Amgen, Inc. (US)

FEATURES
Location/Qualifiers
source
1..4159
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..69
2627..2725
4079..4159
exon
exon
BASE COUNT 1009 a 1037 c 983 g 1110 t 20 others
ORIGIN

Query Match 11.5%; Score 93; DB 6; Length 4159;
Best Local Similarity 100.0%; Pred. No. 4.9e-44;
Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AGGGAAGAGGCGTCTGCTCAAGGCTGCTGAGGAGGAGGAGGCTGCTGCTGCCACCG 156
Db 2634 AGGGAAGAGGCGTCTGCTCAAGGCTGCTGAGGAGGAGGAGGCTGCTGCTGCCACCG 2693

Qy 157 AGTCCCTAGCCCCCACTCAACAAACCTGAAAGG 189
Db 2694 AGTCCCTAGCCCCCACTCAACAAACCTGAAAGG 2726

RESULT 13
AL671190/c
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-464L12 on chromosome 4, complete
sequence.
ACCESSION AL671190.9 GI:21214169
VERSION AL671190.9 GI:21214169
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Griffiths, C.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: clonerequest@sanger.ac.uk
humquery@sanger.ac.uk
On May 25, 2002 this sequence version replaced gi:21213389.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Swi.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
RP11-398M15 is from the library RPC1-11.2 constructed by the group
of Pieter de Jong. For further details see

one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Swi.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
from the RPC1-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACE3.6

FEATURES
Location/Qualifiers
source
1..180176
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-464L12"
/clone_lib="RPC1-23"
BASE COUNT 49850 a 42493 c 40551 g 47282 t
ORIGIN

Query Match 3.3%; Score 27; DB 10; Length 180176;
Best Local Similarity 100.0%; Pred. No. 0.00046;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 757 CAAACAAACAGTAATAATAATAATAA 783
Db 148361 CAAACAAACAGTAATAATAATAATAA 148335

RESULT 14
AL357974/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-398M15 on chromosome 1, complete
sequence.
ACCESSION AL357974
VERSION AL357974.16 GI:14329558
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS Mashreghi-Mohammadi, M.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 8, 2001 this sequence version replaced gi:13751416.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em., EMBL; Swi.,
SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
RP11-398M15 is from the library RPC1-11.2 constructed by the group
of Pieter de Jong. For further details see

<http://www.chori.org/bacpac/home.htm>
VECTOR: pBAC3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-398M15. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true right end of clone RP11-398M15 is at 93165 in this sequence. The true right end of clone RP11-110P20 is at 100 in this sequence.

FEATURES

source	Location/Qualifiers
repeat_region	38062..38665 /note="LIPAL1 repeat: matches 5549..6165 of consensus"
repeat_region	40064..40625 /note="HERVL repeat: matches 4356..4930 of consensus"
repeat_region	40716..40783 /notes="17 copies 4 mer ggag 69% conserved"
repeat_region	41270..41408 /notes="L2 repeat: matches 1273..1413 of consensus"
repeat_region	41409..41691 /note="AluY repeat: matches 1..309 of consensus"
repeat_region	41692..41765 /notes="L2 repeat: matches 1202..1273 of consensus"
repeat_region	41766..42192 /notes="MSTD repeat: matches 1..394 of consensus"
repeat_region	42193..42252 /notes="L2 repeat: matches 1143..1202 of consensus"
repeat_region	43672..43854 /notes="MER5A repeat: matches 1..183 of consensus"
repeat_region	44566..44647 /notes="LIME repeat: matches 5395..5473 of consensus"
repeat_region	45074..45317 /note="MIR repeat: matches 4..259 of consensus"
repeat_region	45463..45656 /notes="MIR repeat: matches 74..262 of consensus"
repeat_region	45929..46050 /notes="L2 repeat: matches 2626..2750 of consensus"
repeat_region	46167..46272 /notes="L2 repeat: matches 2384..2501 of consensus"
repeat_region	47122..47163 /notes="21 copies 2 mer ac 92% conserved"
repeat_region	47124..47163 /note="10 copies 4 mer acac 95% conserved"
repeat_region	47825..48134 /notes="AluSg repeat: matches 1..310 of consensus"
repeat_region	48197..48292 /notes="L2 repeat: matches 2662..2739 of consensus"
repeat_region	48293..48670 /notes="THE1C repeat: matches 1..371 of consensus"
repeat_region	48671..48763 /notes="L2 repeat: matches 2564..2662 of consensus"
repeat_region	48766..48933 /notes="MER69A repeat: matches 5..177 of consensus"
repeat_region	49075..49393 /notes="L2 repeat: matches 69..386 of consensus"
repeat_region	51782..52105 /note="MER44A repeat: matches 3..333 of consensus"
repeat_region	53670..54781 /notes="LIPAL5 repeat: matches 5065..6157 of consensus"
repeat_region	54782..55089 /notes="AluSp repeat: matches 3..310 of consensus"
repeat_region	55090..55691 /note="LIPAL5 repeat: matches 4480..5065 of consensus"
repeat_region	55692..56707 /notes="LIP repeat: matches 2855..3886 of consensus"
repeat_region	56880..56986 /notes="LIM4 repeat: matches 2949..3054 of consensus"
repeat_region	57122..57177 /note="14 copies 4 mer tata 85% conserved"
repeat_region	57743..57880 /notes="L2 repeat: matches 2621..2742 of consensus"
repeat_region	58525..58570 /notes="MSTA repeat: matches 381..426 of consensus"
repeat_region	58573..58866 /note="AluSp repeat: matches 4..307 of consensus"
repeat_region	58867..58942 /notes="19 copies 4 mer ggaa 89% conserved"
repeat_region	58962..59333 /note="MSTA repeat: matches 6..399 of consensus"
repeat_region	59428..59494 /notes="L2 repeat: matches 1951..2023 of consensus"
repeat_region	59495..59717 /notes="MER58A repeat: matches 2..223 of consensus"
repeat_region	59718..59844

repeat_region	843..965 /note="L2 repeat: matches 2605..2740 of consensus"
repeat_region	1154..11770 /db_xref="taxon:9606"
repeat_region	1771..2127 /note="L2 repeat: matches 1841..2413 of consensus"
repeat_region	2128..2628 /note="MLM1A1 repeat: matches 1..365 of consensus"
repeat_region	3949..3990 /note="L2 repeat: matches 1216..1841 of consensus"
repeat_region	7352..7720 /note="LTR40b repeat: matches 421..462 of consensus"
repeat_region	8235..8278 /note="LIP4 repeat: matches 5777..6146 of consensus"
repeat_region	8831..8980 /note="11 copies 4 mer caca 100% conserved"
repeat_region	8837..8980 /note="75 copies 2 mer at 85% conserved"
repeat_region	9022..9093 /note="36 copies 4 mer atat 86% conserved"
repeat_region	9284..9367 /note="L2 repeat: matches 2679..2750 of consensus"
repeat_region	9368..10098 /note="LIM4 repeat: matches 3167..3240 of consensus"
repeat_region	10099..10174 /note="LIPAL3 repeat: matches 5419..6155 of consensus"
repeat_region	10237..12387 /note="LIM4 repeat: matches 3061..3167 of consensus"
repeat_region	12458..13627 /note="LIMEC repeat: matches 1379..3033 of consensus"
repeat_region	14898..21039 /note="LIMEC repeat: matches 102..1288 of consensus"
repeat_region	22297..23011 /note="LIPAL3 repeat: matches 1..6145 of consensus"
repeat_region	23284..23408 /note="MIR repeat: matches 118..248 of consensus"
repeat_region	23442..24061 /note="MER67C repeat: matches 1..709 of consensus"
repeat_region	24234..24297 /note="MER92C repeat: matches 489..552 of consensus"
repeat_region	24564..25039 /note="MER89 repeat: matches 1..559 of consensus"
repeat_region	25623..25911 /note="AluSc repeat: matches 1..301 of consensus"
repeat_region	28551..30612 /note="LIMEC repeat: matches 4265..6328 of consensus"
repeat_region	31945..32066 /note="MSTA repeat: matches 305..426 of consensus"
repeat_region	32574..33996 /note="LIMEC repeat: matches 1317..2374 of consensus"
repeat_region	33997..34097 /note="L2 repeat: matches 2146..2247 of consensus"
repeat_region	34152..34322 /note="L2 repeat: matches 2552..2732 of consensus"
repeat_region	35014..35066 /note="L2 repeat: matches 2657..2709 of consensus"
repeat_region	38011..38055 /note="5S repeat: matches 1..45 of consensus"

Tue May 20 09:21:57 2003

```
/note="L2 repeat: matches 1849. .1951 of consensus"
61894. .62182
/note="AluY repeat: matches 1. .289 of consensus"
62183. .62230
/note="L2 copies 4 mer taaa 93% conserved"
62351. .62659
/note="AluSp repeat: matches 1. .311 of consensus"
62847. .63179
/note="MLT1A1 repeat: matches 1. .365 of consensus"
64094. .64321
/note="L1Mc repeat: matches 1515. .1742 of consensus"
64649. .64953
/note="MLT1E repeat: matches 264. .568 of consensus"

Query Match      3.1% Score 25; DB 9; Length 93165;
Best Local Similarity 100.0%; Pred.No.0.0073;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 763 AACAGTAATAATAATAATAATAA 787
Db 55365 AACAGTAATAATAATAATAATAA 55341

RESULT 15
AC022829/c
LOCUS
DEFINITION Homo sapiens clone RP11-121C1, WORKING DRAFT SEQUENCE, 20 unordered
pieces.
ACCESSION AC022829
VERSION AC022829.3 GI:7249268
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 152245)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-121C1
Unpublished
2 (bases 1 to 152245)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bada,F.,
Boguslavskiy,B., Boukhgaltier,B., Brown,A., Burkett,G., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
Ferrelira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczyk,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivat,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Tjrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
JOURNAL
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT
On Mar 16, 2000 this sequence version replaced gi:6984468.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6137
Center clone name: 121_C.1
----- Summary Statistics
```

```
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 142478 bases at least Q40
Consensus quality: 147185 bases at least Q30
Consensus quality: 148983 bases at least Q20
Insert size: 150000; agarose-fp
Quality coverage: 4.3 in Q20 bases; agarose-fp
Quality coverage: 4.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 20 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1283: contig of 1283 bp in length
* 1284 1383: gap of 100 bp
* 1384 1916: contig of 533 bp in length
* 1917 2016: gap of 100 bp
* 2017 3647: contig of 1631 bp in length
* 3648 3747: gap of 100 bp
* 3748 5481: contig of 1734 bp in length
* 5482 5581: gap of 100 bp
* 5582 6718: contig of 1137 bp in length
* 6719 6818: gap of 100 bp
* 6819 10904: contig of 4086 bp in length
* 10905 11004: gap of 100 bp
* 11005 14393: contig of 3389 bp in length
* 14394 14493: gap of 100 bp
* 14494 18255: contig of 3762 bp in length
* 18256 18355: gap of 100 bp
* 18356 23193: contig of 4838 bp in length
* 23194 23293: gap of 100 bp
* 23294 28506: contig of 5213 bp in length
* 28507 28606: gap of 100 bp
* 28607 36308: contig of 7702 bp in length
* 36309 36408: gap of 100 bp
* 36409 43542: contig of 7134 bp in length
* 43543 43642: gap of 100 bp
* 43643 51047: contig of 7405 bp in length
* 51048 51147: gap of 100 bp
* 51148 61019: contig of 9872 bp in length
* 61020 61119: gap of 100 bp
* 61120 74018: contig of 12899 bp in length
* 74019 74118: gap of 100 bp
* 74119 84938: contig of 10820 bp in length
* 84939 85038: gap of 100 bp
* 85039 100034: contig of 14956 bp in length
* 100035 100134: gap of 100 bp
* 100135 112980: contig of 12846 bp in length
* 112981 113080: gap of 100 bp
* 113081 130270: contig of 17190 bp in length
* 130271 130370: gap of 100 bp
* 130371 152245: contig of 21875 bp in length.
FEATURES
Location/Qualifiers
source
1. .152245
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="RP11-121C1"
/clone_lib="RPCI-11 Human Male BAC"
misc_feature
1. .1283
/note="assembly_fragment"
1384. .1916
/note="assembly_fragment"
clone_end:77
vector_side:left
misc_feature
2017. .3647
/note="assembly_fragment"
misc_feature
3748. .5481
```

	/note="assembly_fragment	
misc_feature	5882..6718	
	/note="assembly_fragment	
	clone_end:Spg	
	vector_side:right"	
misc_feature	6819..10904	
	/note="assembly_fragment"	
misc_feature	11005..14393	
	/note="assembly_fragment"	
misc_feature	14494..18255	
	/note="assembly_fragment"	
misc_feature	18356..23193	
	/note="assembly_fragment"	
misc_feature	23294..28506	
	/note="assembly_fragment"	
misc_feature	28607..36308	
	/note="assembly_fragment"	
misc_feature	36409..43542	
	/note="assembly_fragment"	
misc_feature	43643..51047	
	/note="assembly_fragment"	
misc_feature	51148..61019	
	/note="assembly_fragment"	
misc_feature	61120..74018	
	/note="assembly_fragment"	
misc_feature	74119..84938	
	/note="assembly_fragment"	
misc_feature	85039..100034	
	/note="assembly_fragment"	
misc_feature	100135..112980	
	/note="assembly_fragment"	
misc_feature	113081..130270	
	/note="assembly_fragment"	
misc_feature	130371..152245	
	/note="assembly_fragment"	
BASE COUNT	47920 a 28046 c 27243 g 47134	
ORIGIN		

```

Query Match      3.1%; Score 25; DB 2; Length 152245;
Best Local Similarity 100.0%; Pred. No. 0.0075;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 770 AATAAATAAAATAAACTCAATG 794
      |||||
Db 43991 AATAAATAAAATAAACTCAATG 43967

```

Search completed: May 17, 2003, 16:50:03
Job time : 3076 secs

PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 10-NOV-2000; 2000WO-US30873.
XX (GETH) GENENTECH INC.
XX Baker KP, Bereini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI: 2001-408281/43.
DR P-PSDB; AAU12246.
XX Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical -
XX Claim 3; Fig 149; 813pp; English.
XX AAS21244-AAS21518 encode for novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIa. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
SQ Sequence 804 BP; 199 A; 263 C; 176 G; 166 T; 0 other;
Query Match 97.0%; Score 782; DB 22; Length 804;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTCTGCT 69
Db |||||
Qy 23 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCTCTTCCAGGCTGCTCTGTATCTCTGCT 82
Db |||||
Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGGAAGGCGTCTGCGAAGGCTGCTGTCAGG 129
Db |||||
Qy 83 TCTCTGCTTCTCCATCTTCTCCACAGAGGGAAGGCGTCTGCGAAGGCTGCTGTCAGG 142
Db |||||
Qy 130 CAGGAGAACCAAGGCTCTGTCGCCACCGAGTCCCTAGCCGCCCACTCAACAAACCTGAAAGG 189
Db |||||
Qy 143 CAGGAGAACCAAGGCTCTGTCGCCACCGAGTCCCTAGCCGCCCACTCAACAAACCTGAAAGG 202
Db |||||
Qy 190 ACATCATGTGAGGCTCTGTAAACCATCAAGCTTGAGCCAGAGCCCCGCCCTTTGGTGGT 249
Db |||||
Qy 203 ACATCATGTGAGGCTCTGTAAACCATCAAGCTTGAGCCAGAGCCCCGCCCTTTGGTGGT 262
Db |||||
Qy 250 GCCTGGGGCACTCCACAGGTTAGCACTCCCAAGCAAGACTCCAGACGGGAGAAC 309
Db |||||
Qy 263 GCCTGGGGCACTCCACAGGTTAGCACTCCCAAGCAAGACTCCAGACGGGAGAAC 322
Db |||||
Qy 310 TCATGCTGGCACTCCAGGTTAGCACTCCCAAGCAAGACTCCAGACGGGAGAAC 369
Db |||||
Qy 323 TCATGCTGGCACTCCAGGTTAGCACTCCCAAGCAAGACTCCAGACGGGAGAAC 382
Db |||||
Qy 370 GTGAGCTCCAATGTTGGAGGCTTCATCTCGGGTGCAGGACCTCGGGAAGTTCCAGA 429
Db |||||

Db 383 GTGAGCTCCAATGTTGGAGGCTTCTATCTCGGGCTGCAAGACCCCTGGGAAAGTTCCAGA 442
Qy 430 ACTCCACGCTCTTGTCTCTCAATTTGTGCCATCAACTTTTTCAGAGCTATCATGAGCCAACTCA 489
Db |||||
Db 443 ACTCCACGCTCTTGTCTCTCAATTTGTGCCATCAACTTTTTCAGAGCTATCATGAGCCAACTCA 502
Qy 490 CCCACAGGGGCTCAGTTCGGCCACCATGTGGGCTCTCTCCAGTGCMAACCCAGGAGCATTC 549
Db 503 CCCACAGGGGCTCAGTTCGGCCACCATGTGGGCTCTCTCCAGTGCMAACCCAGGAGCATTC 562
Qy 550 ACCATGACCGGTTCACAGCTCAAAATCCAGAGACCATCAATCCCTGCTAGAGTGCAGGGTGG 609
Db 563 ACATGACCGGTTCACAGCTCAAAATCCAGAGACCATCAATCCCTGCTAGAGTGCAGGGTGG 622
Qy 610 CAAGCACCCCAAGGGTGTGCTGACCAAGACTCCAGAGTCTCTCCATCTTTCAGGTCATTC 669
Db 623 CAAGCACCCCAAGGGTGTGCTGACCAAGACTCCAGAGTCTCTCCATCTTTCAGGTCATTC 682
Qy 670 GCTTCTGCGCATTTAACTACCAAGCATCCAGTGTGCTCCCAAGGAATCCCTTCTAGCCTCC 729
Db 683 GCTTCTGCGCATTTAACTACCAAGCATCCAGTGTGCTCCCAAGGAATCCCTTCTAGCCTCC 742
Qy 730 TGACATGAGTCTGCTGGAAGAGCATCCAAACAAAGTAATAATAATAATAATAATAA 789
Db 743 TGACATGAGTCTGCTGGAAGAGCATCCAAACAAAGTAATAATAATAATAATAATAA 802
Qy 790 CA 791
Db |||
Db 803 CA 804
RESULT 3
AAF97898
ID AAF97898 standard; cDNA; 908 BP.
XX
AC AAF97898;
XX
DT 01-JUN-2001 (first entry)
XX
DE Human secreted protein cDNA, SEQ ID NO: 25.
XX
KW Human; secreted protein; immunomodulatory; antisclerotic;
KW dermatological; antiinflammatory; anti-HIV; cytostatic; cardiant;
KW vascular; anti-angiogenic; ophthalmological; neuroprotectant;
KW neotrophic; anticonvulsant; antialzheimers; antiparkinsonian;
KW antimicrobial; vulneryary; vaccine; gene therapy; cancer;
KW protein coordinate data; infection; ss.
XX
OS Homo sapiens.
XX
PN WO200121658-A1.
XX
PD 29-MAR-2001.
XX
PF 22-SEP-2000; 2000WO-US26013.
XX
PR 24-SEP-1999; 99US-0155709.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ni J, Baker KP, Birse CE, Ebner R, Fiscella M, Komatsoulis GA;
PI Lafleur DW, Moore PA, Olsen HS, Rosen CA, Ruben SA, Soppet DR;
PI Young PE, Wei P, Florence KA;
XX
DR WPI: 2001-235311/24.
XX
PT Nucleic acids encoding 32 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -
XX
PS Claim 1; Page 730; 890pp; English.
XX
CC The present sequence encodes one of 32 novel human secreted polypeptides.

CC The nucleic acid molecules and polypeptides they encode may be used in
CC the prevention, diagnosis and treatment of diseases such as
CC immune disorders (e.g. multiple sclerosis, systemic lupus erythematosus
CC and human immuno-deficiency virus (HIV) infections), hyperproliferative
CC disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
CC (e.g. Schmitz syndrome, Chaga's cardiomyopathy and coronary
CC arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC neovascularisation and diabetic retinopathy), neurological disorders
CC (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
CC infectious diseases and/or for promoting wound healing, regeneration
CC and/or chemotaxis. The nucleic acid molecules may be used to produce the
CC secreted polypeptides. They may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acid
CC sequences in samples. The polypeptides may be used as antigens in the
CC production of antibodies and in assays to identify modulators of
CC their expression and activity.

XX
SQ Sequence 908 BP; 242 A; 290 C; 197 G; 175 T; 4 other;

Query Match 90.8%; Score 732; DB 22; Length 908;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 782; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 10 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCCCTTTCCAGCCTGCTCTGTATCTCTCT 69
DB 66 GAAATCTGCTTCTCACCATGAGGCTTCTAGTCCCTTTCCAGCCTGCTCTGTATCTCTCT 125
QY 70 TCTCTGCTTCTCCTTCTCACAAGAGGAGGGCTCTGCCAAGGCTGGTCAGG 129
DB 126 TCTCTGCTTCTCCTTCTCACAAGAGGAGGGCTCTGCCAAGGCTGGTCAGG 185
QY 130 CAGGAGAACAGGCTCTGCTGCCACCGAGTCCCTAGCCCACTCAACAAACCTGAAAGG 189
DB 186 CAGGAGAACAGGCTCTGCTGCCACCGAGTCCCTAGCCCACTCAACAAACCTGAAAGG 245
QY 190 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCGAGGCCCCCTTTGGGTGGT 249
DB 246 ACATCATGTGAGGCTCTGTAAACCATGCAAGCTTGAGCGAGGCCCCCTTTGGGTGGT 305
QY 250 GCCTGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGGGAGAAC 309
DB 306 GCCTGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGGGAGAAC 365
QY 310 TCATCGCTGGCACTCCAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGGGAGAAC 369
DB 366 TCATCGCTGGCACTCCAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGGGAGAAC 425
QY 370 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGA 429
DB 426 GTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGA 485
QY 430 ACTCCAGCTCCTTGTCTCAATGTGCCATCACTTTTCAGAGTATCATGAGCCCACTCA 489
DB 486 ACTCCAGCTCCTTGTCTCAATGTGCCATCACTTTTCAGAGTATCATGAGCCCACTCA 545
QY 490 CCCACAGGGCTCAGTCGCAACCATGTGGGCTCTCCAGTGCAACCAACGAGCATTC 549
DB 546 CCCACAGGGCTCAGTCGCAACCATGTGGGCTCTCCAGTGCAACCAACGAGCATTC 605
QY 550 ACCATGACCGGTTCACAGCTTACAAATCCAGAGACCAATCAATCTGTAGAGTGCAGGGTGG 609
DB 606 ACCATGACCGGTTCACAGCTTACAAATCCAGAGACCAATCAATCTGTAGTGCAGGGTGG 665
QY 610 CAGACACCAAGGGTGGCTGACCAAGACTCGAGAGTCTCCCTCATCTTCAGGTCCATTCA 669
DB 666 CAGACACCAAGGGTGGCTGACCAAGACTCGAGAGTCTCCCTCATCTTCAGGTCCATTCA 725
QY 670 GCCTCTGGCATTTAACTTACAGCATCCAGTGTGTCCTCCCAAGGAATCTCTTACGCTCC 729
DB 726 GCCTCTGGCATTTAACTTACAGCATCCAGTGTGTCCTCCCAAGGAATCTCTTACGCTCC 785
QY 730 TGACATGAGTCTGTGGAAAGAGCATCAAAACAAAGTAAATAAATAAATAAATAA 789

DB 786 TGACATGAGTCTGTGGAAAGAGCATCAAAACAAAGTAATAAATAAATAAATAA 845
QY 790 CAA 792
DB 846 CAA 848
RESULT 4
AA72224/c
ID AAA72224 standard; cDNA; 801 BP.
XX
AC AAA72224;
XX
DT 06-DEC-2000 (first entry)
XX
DE Human CASB611 cDNA.
XX
KW Human; CASB611; colon-specific expression; expressed sequence tag;
KW EST; colon cancer; tumour; autoimmune disease; diagnosis;
KW disease susceptibility; prophylaxis; genetic vaccine; gene therapy; ss.
OS Homo sapiens.
XX
PN WO200043509-A2.
XX
PD 27-JUL-2000.
XX
PF 17-JAN-2000; 2000WO-EP00346.
XX
PR 19-JAN-1999; 99GB-0001078.
PR 29-JAN-1999; 99GB-0002090.
PR 01-FEB-1999; 99GB-0002163.
PR 01-FEB-1999; 99GB-0002168.
PR 01-FEB-1999; 99GB-0002169.
PR 07-APR-1999; 99GB-0007901.
XX
PA (SMIK) SMITHLINE BERCHAM BIOLOGICALS.
XX
PI Vinals-Bassols C;
XX
XX WPI; 2000-482912/42.
XX
PT New isolated polynucleotide useful for diagnosis and/or treatment of
PT colon cancer and autoimmune disease -
XX
PS Claim 3; Page 34; 41pp; English.
XX
CC This sequence represents human CASB611 cDNA. This gene exhibits
CC colon-specific expression and is highly expressed in the rectum.
CC The invention relates to human CASB cDNA sequences CASB611, CASB500,
CC CASB501, CASB502, CASB505 and CASB507 (AA72224-A72229, respectively)
CC and also to human CASB partial cDNA sequences (AA72230-A72233) derived
CC from expressed sequence tags (ESTs). Expression of the human CASB genes
CC (with the exception of CASB611) is associated with colon tumours, and
CC the encoded proteins (sequences not given in the specification)
CC represent colon tumour-associated antigens. The cDNA sequences may be
CC used in diagnosing the presence or a susceptibility to a disease related
CC to the presence, expression or activity of CASB genes. Such diseases
CC include autoimmune diseases and especially colon cancer. The nucleic
CC acid sequences may also be used in genetic vaccines for the prophylaxis
CC or therapeutic treatment of colon cancer and autoimmune diseases.
XX
SQ Sequence 801 BP; 165 A; 181 C; 255 G; 200 T; 0 other;
Query Match 66.6%; Score 537; DB 21; Length 801;
Best Local Similarity 100.0%; Pred. No. 2e-246;
Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 261 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACGGGAGACCTCATGCTGGC 320
DB 537 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACGGGAGACCTCATGCTGGC 478
QY 321 ACCTGAGGTACCCAGCAGCCTCTGTCTCCCTTTCAGCCTTCACAGCAGTGAGCTCAA 380

Db 477 ACCTGAGGTACCCAGCAGCCTCTGTCTCCCTTTTTCAGCCTTTTTCAGCAGTGTAGCTGCAA 418
Qy 381 TGTTCGAGGGCTTCATCTCGGGCTGCAAGGACCTTGGGAAAGTTCCAGAACTCCACGTCC 440
Db 417 TGTTCGAGGGCTTCATCTCGGGCTGCAAGGACCTTGGGAAAGTTCCAGAACTCCACGTCC 358
Qy 441 TTGTCTCAATTTGTGCATCAACTTTTCAGAGCTATCATGAGCAACTCACCCACAGGGC 500
Db 357 TTGTCTCAATTTGTGCATCAACTTTTCAGAGCTATCATGAGCAACTCACCCACAGGGC 298
Qy 501 CTCAGTCCCAACCATGTGGGCTCTCCAGTGCAGGCAACCCAGGCAATTCACCATGACCGG 560
Db 297 CTCAGTCCCAACCATGTGGGCTCTCCAGTGCAGGCAACCCAGGCAATTCACCATGACCGG 238
Qy 561 TCACAGTACAAATCCAGAGCACCATCAATCTCTGTAGAGTGCAGGCTGGCAAGCACCAC 620
Db 237 TCACAGTACAAATCCAGAGCACCATCAATCTCTGTAGAGTGCAGGCTGGCAAGCACCAC 178
Qy 621 GGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGGCTCTCTGCA 680
Db 177 GGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGGCTCTCTGCA 118
Qy 681 TTTAACTACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGGCTCTCTGCA 740
Db 117 TTTAACTACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGGCTCTCTGCA 58
Qy 741 TGCTGGAAGAGCATCCAAACCAAGTAAATAAATAAATAAATAAATAAATAAATAAATGCG 797
Db 57 TGCTGGAAGAGCATCCAAACCAAGTAAATAAATAAATAAATAAATAAATAAATAAATGCG 1

RESULT 5
AAA72230/C
ID AAA72230 standard; cDNA; 797 BP.
XX AC AAA72230;
XX DT 06-DEC-2000 (first entry)
XX DE Human CASB gene partial cDNA, SEQ ID NO:7.
XX KW Human; CASB gene; overexpression; colon tumour-associated antigen;
KW expressed sequence tag; EST; colon cancer; tumour; autoimmune disease;
KW diagnosis; disease susceptibility; prophylaxis; genetic vaccine;
KW gene therapy; ss.
XX OS Homo sapiens.
XX FN WO200043509-A2.
XX PD 27-JUL-2000.
XX PF 17-JAN-2000; 2000WO-EP00346.
XX PR 19-JAN-1999; 99GB-0001078.
PR 29-JAN-1999; 99GB-0002090.
PR 01-FEB-1999; 99GB-0002163.
PR 01-FEB-1999; 99GB-0002168.
PR 01-FEB-1999; 99GB-0002169.
PR 07-APR-1999; 99GB-0007901.
XX PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX PI Vinale-Bassols C;
XX DR WPI; 2000-482912/42.
XX PT New isolated polynucleotide useful for diagnosis and/or treatment of
PT colon cancer and autoimmune disease -
XX FS Disclosure; Page 35-36; 41pp; English.
XX

CC Sequences AAA72230-A72235 represent human CASB gene partial cDNA
CC sequences which are derived from expressed sequence tags (ESTs). The
CC invention relates to human CASB cDNA sequences CASB611, CASB500,
CC CASB501, CASB502, CASB505 and CASB507 (AAA72224-A72229, respectively) and
CC also to these human CASB partial cDNA sequences. Expression of the human
CC CASB genes (with the exception of CASB611) is associated with colon
CC tumours, and the encoded proteins (sequences not given in the
CC specification) represent colon tumour-associated antigens. The cDNA
CC sequences may be used in diagnosing the presence or a susceptibility to
CC a disease related to the presence, expression or activity of CASB genes.
CC Such diseases include autoimmune diseases and especially colon cancer.
CC The nucleic acid sequences may also be used in genetic vaccines for the
CC prophylaxis or therapeutic treatment of colon cancer and autoimmune
CC diseases.
XX
SQ Sequence 797 BP; 164 A; 180 C; 254 G; 199 T; 0 other;
Query Match 61.8%; Score 498; DB 21; Length 797;
Best Local Similarity 100.0%; Pred. No. 8.5e-228;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 261 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAACCTCATGTGCTGGC 320
Db 533 TCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAACCTCATGTGCTGGC 474
Qy 321 ACCTGAGTACCCAGCAGCCTCTGTCTCCCTTTTCAGCCTTCACAGCAGTGTGCA 380
Db 473 ACCTGAGTACCCAGCAGCCTCTGTCTCCCTTTTCAGCCTTCACAGCAGTGTGCA 414
Qy 381 TGTTCGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCAGTCC 440
Db 413 TGTTCGAGGGCTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCCAGAACTCCAGTCC 354
Qy 441 TTGTCTCAATTTGTGCATCAACTTTTCAGAGCTATCATGAGCACAACCTCACCCACAGGGC 500
Db 353 TTGTCTCAATTTGTGCATCAACTTTTCAGAGCTATCATGAGCACAACCTCACCCACAGGGC 294
Qy 501 CTCAGTCCCAACCATGTGGGCTCTTCCAGTGCAGGCAACCCAGGCAATTCACCATGACCGG 560
Db 293 CTCAGTCCCAACCATGTGGGCTCTTCCAGTGCAGGCAACCCAGGCAATTCACCATGACCGG 234
Qy 561 TCACAGTACAAATCCAGAGCACCATCAATCTCTGTAGAGTGCAGGCTGGCAAGCACCAC 620
Db 233 TCACAGTACAAATCCAGAGCACCATCAATCTCTGTAGAGTGCAGGCTGGCAAGCACCAC 174
Qy 621 GGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGGCTCTCTGCA 680
Db 173 GGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGGCTCTCTGCA 114
Qy 681 TTTAACTACCAAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCTCTGACATGAGTC 740
Db 113 TTTAACTACCAAGCATCCAGTGGTCCCAAGGAATCCCTTCTAGCCTCTGACATGAGTC 54
Qy 741 TGCTGGAAGAGCATCCCA 758
Db 53 TGCTGGAAGAGCATCCCA 36
RESULT 6
ABK27569
ID ABK27569 standard; cDNA; 485 BP.
XX AC ABK27569;
XX AC ABK27569;
XX DT 09-APR-2002 (first entry)
XX DE Human colon cancer expressed sequence tag, Seq ID no 6.
XX KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;
KW expressed sequence tag.
XX OS Homo sapiens.
XX

PN WO200196390-A2.
XX 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18557.
XX
XX 09-JUN-2000; 2000US-210821P.
XX PR 18-DEC-2000; 2000US-256571P.
XX PR 10-MAY-2001; 2001US-290240P.
XX
PA (CORI-) CORIXA CORP.
XX
XX Jiang Y, Hepler WT, Clapper JD, Wang A, Secretist H;
PI WPI; 2002-139708/18.
XX
XX Novel isolated polynucleotide encoding a polypeptide comprising a
PT portion of colon tumour protein, useful for detection, diagnosis and
PT therapy of human colon cancer
XX
XX Claim 1; Page 151-152; 220pp; English.
XX
XX The invention relates to an isolated polynucleotide (I) encoding a
CC polypeptide (II) comprising at least a portion of a colon tumour
CC protein. (I), (II) and antibody (III) to (II) are useful for determining
CC the presence of a cancer in a patient. (I), (II) or antigen presenting
CC cells expressing (I) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein, by contacting T cells with (I), (II) or
CC antigen-presenting cells that express (I), under conditions and for a
CC time sufficient to permit the stimulation and/or expansion of T cells.
CC (I), (II), or antigen presenting cells that express (II) are useful for
CC treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells
CC isolated from a patient with (I), (II) or antigen presenting cells that
CC express (II), such that T cells proliferate, and administering to the
CC patient an effective amount of the proliferated T cells, thus inhibiting
CC the development of a cancer in the patient. (I) or (II) is useful in
CC vaccines and pharmaceutical compositions for prevention and treatment
CC of colon malignancies and for the diagnosis and monitoring of such
CC cancers. (I) (II) or (III) is useful for detection, diagnosis and/or
CC therapy of human colon cancer. (I) is useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of
CC ribozyme molecules for inhibiting expression of (II) in tumour cells.
CC ABK27564-ABK27807 represent novel human colon cancer coding
CC sequences and primers of the invention.
XX
SQ Sequence 485 BP; 109 A; 155 C; 124 G; 97 T; 0 other;
Query Match 40.1%; Score 323; DB 24; Length 485;
Best Local Similarity 100.0%; Pred. No. 3.3e-144;
Matches 323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10 GAAATCTGCCTTCCACCATGAGGCTTCTAGTCTCTTCCAGCCTGCTGTATCTGCT 69
Db 60 GAAATCTGCCTTCCACCATGAGGCTTCTAGTCTCTTCCAGCCTGCTGTATCTGCT 119
Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGAGGCTCTGCTGCAAGGCTGTGTCAGG 129
Db 120 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGAGGCTCTGCTGCAAGGCTGTGTCAGG 179
Qy 130 CAGGAGAACCCAGGCTCTGCTGCCACCGAGTCCCTAGGCCCACTCAACAAACCTGAAAGG 189
Db 180 CAGGAGAACCCAGGCTCTGCTGCCACCGAGTCCCTAGGCCCACTCAACAAACCTGAAAGG 239
Qy 190 ACATCATGTGAGGCTCTGTAACCATGACGCTTGGAGCAGAGCCCGCCCTTTGGGTGGT 249
Db 240 ACATCATGTGAGGCTCTGTAACCATGACGCTTGGAGCAGAGCCCGCCCTTTGGGTGGT 299
Qy 250 GCCTGGGGCACTCCACAGGCTGTAGCACTCCCAAGAGCAAGACTCCAGACAGCGGAGAAC 309
Db 300 GCCTGGGGCACTCCACAGGCTGTAGCACTCCCAAGAGCAAGACTCCAGACAGCGGAGAAC 359
Qy 310 TCATGCTGGCACTTGGAGTACC 332

Db 360 TCATGCTGGCACCTGAGGTACC 382
RESULT 7
ABL37082/c
ID ABL37082 standard; cDNA; 382 BP.
XX
AC ABL37082;
XX
DT 08-APR-2002 (first entry)
XX
XX Human colon tumour antigen polynucleotide SEQ ID NO:671.
DE Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
XX colon tumour metastatic antigen; diagnosis; gene; ss.
KW
XX Homo sapiens.
OS
XX WO200196388-A2.
PN
XX 20-DEC-2001.
XX
XX 08-JUN-2001; 2001WO-US18557.
XX
XX 09-JUN-2000; 2000US-210899P.
XX PR 20-FEB-2001; 2001US-270216P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Harlocker SL, Secretist H;
PI WPI; 2002-114514/15.
XX
XX Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient
XX
XX Claim 1; SEQ ID 671; 105pp; English.
XX
XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (I)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (I) have cytostatic activity and can be used in vaccine
CC production. (I) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (I) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (I)
CC can be used in the diagnosis of a colon tumour.
XX
SQ Sequence 382 BP; 76 A; 94 C; 122 G; 86 T; 4 other;
Query Match 37.1%; Score 299; DB 24; Length 382;
Best Local Similarity 100.0%; Pred. No. 9.7e-133;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 10 GAAATCTGCCTTCTCACCATGAGGCTTCTAGTCTCTTCCAGCCTGCTGTATCTGCTGCT 69
Db 320 GAAATCTGCCTTCTCACCATGAGGCTTCTAGTCTCTTCCAGCCTGCTGTATCTGCTGCT 261
Qy 70 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGAGGCTCTGCTGCAAGGCTGTGTCAGG 129
Db 260 TCTCTGCTTCTCCATCTTCTCCACAGAGGAGGAGGCTCTGCTGCAAGGCTGTGTCAGG 201
Qy 130 CAGGAGAACCCAGGCTCTGCTGCCACCGAGTCCCTAGGCCCACTCAACAAACCTGAAAGG 189
Db 200 CAGGAGAACCCAGGCTCTGCTGCCACCGAGTCCCTAGGCCCACTCAACAAACCTGAAAGG 141
Qy 190 ACATCATGTGAGGCTCTGTAACCATGACGCTTGGAGCAGAGCCCGCCCTTTGGGTGGT 249
Db 140 ACATCATGTGAGGCTCTGTAACCATGACGCTTGGAGCAGAGCCCGCCCTTTGGGTGGT 81
Qy 250 GCCTGGGGCACTCCACAGGCTGTAGCACTCCCAAGAGCAAGACTCCAGACAGCGGAGAAC 308

Db 80 GCCTGGGGCACTCCACAGGTGTAGCACTCCCAAGCAAGACTCCAGACAGCGGAGAAC 22
RESULT 8
ID AAS39569 standard; cDNA; 487 BP.
XX AAS39569;
AC AAS39569;
XX
DT 17-DEC-2001 (first entry)
XX
DE cDNA encoding novel human colon associated polypeptide #22.
XX
KW Human; colon cancer; congenital abnormality; infection; colitis;
KW inflammatory bowel disease; IBD; neoplastic disorder; gene therapy;
KW intestinal inflammatory disorder; malabsorption syndrome; gastric;
KW sigmoid disease; antibacterial; antiviral; antiinflammatory;
KW cytostatic; ss.
XX
OS Homo sapiens.
XX
FN WO200155302-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01240.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214986.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.


```
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
```

```
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-502630/55.
XX P-PSDB; AAM93135.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases -
XX
XX Claim 1; SEQ ID NO 1224; 986pp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a cDNA encoding a digestive
XX system antigen of the invention.
XX
XX Sequence 487 BP; 135 A; 156 C; 100 G; 95 T; 1 other;
SQ
Query Match 28.7%; Score 231; DB 22; Length 487;
Best Local Similarity 99.6%; Pred. No. 2.9e-100;
Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 375 CTGCAATGTTGGAGGGTTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCAGAACTCC 434
Db 69 CTGCAATGTTGGAGGGTTTCATCTCGGGCTGCAAGGACCTGGGAAAGTTCAGAACTCC 128
QY 435 ACGTCCTTGCTCAATTGTGCCATCACTTCAGAGCTATCATGAGCCAACTCACCCCA 494
Db 129 ACGTCCTTGCTCAATTGTGCCATCACTTCAGAGCTATCATGAGCCAACTCACCCCA 188
QY 495 CAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCACCAACCCAGCATTCCACCAT 554
Db 189 CAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCACCAACCCAGCATTCCACCAT 248
QY 555 GACCGGTCACAGCTACAAATCCAGAGACCATCAATCTCTGTAGAGTCAGGGTGGCAAGC 614
Db 249 GACCGGTCACAGCTACAAATCCAGAGACCATCAATCTCTGTAGAGTCAGGGTGGCAAGC 308
QY 615 ACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCT 656
Db 309 ACCCAAGGGTGGCTGACCAAGACTGCAGAGTCTCTCCATCT 350
RESULT 10
AAS02026
ID AAS02026 standard; cDNA; 227 BP.
XX
XX AAS02026;
XX
XX 16-JUL-2001 (first entry)
XX
XX Bladder cancer-associated sequence, TCC96CT.
XX
```

KW Bladder; cancer; transitional cell carcinoma; diagnostic; TCC96C7; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH polyA_site 218
FT /*tag= a
XX WO200122864-A2.
XX 05-APR-2001.
XX 27-SEP-2000; 2000WO-US41005.
XX 27-SEP-1999; 99US-0156153.
XX (QUAR-) QUARK BIOTECH INC.
PA (KOHN/) KOHN K I.
XX Feinstein E, Mor O;
XX WPI; 2001-258076/26.
XX New polynucleotide sequences upregulated in bladder cancer for
PT diagnosing bladder cancer and inhibition of expression is useful for
PT treating and regulating bladder cancer-associated pathologies
XX Claim 4; Page 59; 64pp; English.
XX The sequence represents the coding sequence of bladder cancer-associated
CC sequence, transitional cell carcinoma clone, TCC96C7. The sequence is
CC upregulated in bladder cancer and its expression is indicative of bladder
CC cancer. The sequence can be used as a marker, and can be used for
CC diagnosing bladder cancer. Antibodies, ribozymes, antisense
CC oligonucleotides or a dominant negative peptide directed against the
CC sequence are useful for regulating bladder cancer-associated pathologies
XX in a patient.
XX Sequence 227 BP; 76 A; 63 C; 43 G; 45 T; 0 other;
XX
Query Match 27.3%; Score 220; DB 22; Length 227;
Best Local Similarity 100.0%; Pred. No. 5.3e-95;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 573 ATCCAGACCATCATCTCTAGAGTGCAGGTGGCAAGCACCAGGCTGCTGACC 632
Db 1 ATCCAGACCATCATCTCTAGAGTGCAGGTGGCAAGCACCAGGCTGCTGACC 60
Qy 633 AAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGCTCCTGGCATTTAACTACCAG 692
Db 61 AAGACTGCAGAGTCTCTCCATCTTCAGGTCCATTCAGCTCCTGGCATTTAACTACCAG 120
Qy 693 CATCAGTGGTCCCCAAGGAATCCCTTCTAGCTTCTGACATGAGTCTGTGGAAGAG 752
Db 121 CATCAGTGGTCCCCAAGGAATCCCTTCTAGCTTCTGACATGAGTCTGTGGAAGAG 180
Qy 753 CATCCAAACAACAGTAATAATAATAATAATAATAATAATAATAATAATACTCAA 792
Db 181 CATCCAAACAACAGTAATAATAATAATAATAATAATAATAATAATAATACTCAA 220
RESULT 11
ABN94440/C
ID ABN94440 standard; DNA; 366 BP.
XX AC ABN94440;
XX 13-AUG-2002 (first entry)
XX Gene #938 used to diagnose liver cancer.
XX Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW metastatic liver tumour; cytostatic; expression profile; disease state;
KW

KW disease progression; drug toxicity; drug efficacy; drug metabolism.
XX Homo sapiens.
XX WO200229103-A2.
XX 11-APR-2002.
XX 02-OCT-2001; 2001WO-US30589.
XX 02-OCT-2000; 2000US-237054P.
XX (GENE-) GENE LOGIC INC.
XX Horne D, Alvares C, Peres-Da-Silva S, Vockley JG;
XX WPI; 2002-426119/45.
XX Diagnosing and detecting the progression of liver cancer,
PT hepatocellular carcinoma or metastatic liver tumor in a patient,
PT involves detecting the level of expression of two or more genes in a
PT liver tissue sample
XX Claim 1; SEQ ID NO 938; 298pp; English.
XX The invention relates to a novel method for diagnosing and detecting the
CC progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC tumour in a patient, and differentiating metastatic liver cancer from
CC hepatocellular carcinoma in a patient, involving detecting the level of
CC expression of two or more genes represented in ABN93503-ABN97455 in a
CC tissue sample. The method of the invention has hepatotropic, and
CC cytostatic activity. The method is useful for diagnosing and detecting
CC the progression of liver cancer, hepatocellular carcinoma and metastatic
CC liver carcinoma in a patient. The method is useful for identifying
CC expression profiles which serve as useful diagnostic markers as well as
CC markers that can be used to monitor disease states, disease progression,
CC drug toxicity, drug efficacy and drug metabolism.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 366 BP; 77 A; 80 C; 119 G; 90 T; 0 other;
XX
Query Match 24.1%; Score 196; DB 24; Length 366;
Best Local Similarity 99.6%; Pred. No. 1.5e-83;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 357 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGTGCAGGACCCCTG 416
Db 366 AGCCTTCACAGCAGTGAGCTGCAATGTTGGAGGGCTTCATCTCGGGTGCAGGACCCCTG 307
Qy 417 GGAAAGTTCAGAACTCCAGTCCCTGCTCTCAATTGTCATCACTTCAGAGCTATCA 476
Db 306 GGAAAGTTCAGAACTCCAGTCCCTGCTCTCAATTGTCATCACTTCAGAGCTATCA 247
Qy 477 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCAGAAC 536
Db 246 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCAGAAC 187
Qy 537 CACCGAGCATTTCCACCATGACCGGTTCAGTCAAGTACAAATCCAGAGACCATCAATCCTGCTA 596
Db 186 CACCGAGCATTTCCACCATGACCGGTTCAGTCAAGTACAAATCCAGAGACCATCAATCCTGCTA 127
Qy 597 GAGTGCA 603
Db 126 GAGTGCA 120
RESULT 12
ABL62262/C
ID ABL62262 standard; DNA; 366 BP.
XX AC ABL62262;
XX

XX DT 15-MAY-2002 (first entry)
XX DE Colon adenocarcinoma related gene sequence SEQ ID NO:599.
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX OS Homo sapiens.
XX PN WO200194629-A2.
XX PD 13-DEC-2001.
XX PF 30-MAY-2001; 2001WO-US10838.
XX PR 05-JUN-2000; 2000US-209473P.
XX PR 05-JUN-2000; 2000US-209531P.
XX PR 18-SEP-2000; 2000US-233133P.
XX PR 18-SEP-2000; 2000US-233611P.
XX PR 20-SEP-2000; 2000US-234009P.
XX PR 20-SEP-2000; 2000US-234034P.
XX PR 20-SEP-2000; 2000US-234052P.
XX PR 22-SEP-2000; 2000US-234509P.
XX PR 22-SEP-2000; 2000US-234567P.
XX PR 25-SEP-2000; 2000US-234923P.
XX PR 25-SEP-2000; 2000US-234924P.
XX PR 25-SEP-2000; 2000US-235077P.
XX PR 25-SEP-2000; 2000US-235082P.
XX PR 25-SEP-2000; 2000US-235134P.
XX PR 25-SEP-2000; 2000US-235280P.
XX PR 26-SEP-2000; 2000US-235637P.
XX PR 26-SEP-2000; 2000US-235638P.
XX PR 27-SEP-2000; 2000US-235711P.
XX PR 27-SEP-2000; 2000US-235720P.
XX PR 27-SEP-2000; 2000US-235840P.
XX PR 27-SEP-2000; 2000US-235863P.
XX PR 28-SEP-2000; 2000US-236028P.
XX PR 28-SEP-2000; 2000US-236032P.
XX PR 28-SEP-2000; 2000US-236033P.
XX PR 28-SEP-2000; 2000US-236034P.
XX PR 28-SEP-2000; 2000US-236109P.
XX PR 29-SEP-2000; 2000US-236111P.
XX PR 29-SEP-2000; 2000US-236842P.
XX PR 29-SEP-2000; 2000US-236891P.
XX PR 02-OCT-2000; 2000US-237172P.
XX PR 02-OCT-2000; 2000US-237173P.
XX PR 02-OCT-2000; 2000US-237278P.
XX PR 02-OCT-2000; 2000US-237294P.
XX PR 02-OCT-2000; 2000US-237295P.
XX PR 02-OCT-2000; 2000US-237316P.
XX PR 03-OCT-2000; 2000US-237425P.
XX PR 03-OCT-2000; 2000US-237598P.
XX PR 03-OCT-2000; 2000US-237604P.
XX PR 03-OCT-2000; 2000US-237606P.
XX PR 03-OCT-2000; 2000US-237608P.
XX PR 01-NOV-2000; 2000US-244867P.
XX PR 01-NOV-2000; 2000US-245084P.
XX PA (AVAL-) AVALON PHARM.
XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX WPI; 2002-188264/24.
XX DR Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set
XX Claim 1; SEQ ID 599; 44pp; English.

XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX SQ Sequence 366 BP; 77 A; 80 C; 119 G; 90 T; 0 other;
Query Match 24.3%; Score 196; DB 24; Length 366;
Best Local Similarity 99.6%; Pred. No. 1.5e-83;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 357 AGCCTTCACAGCAGTGAGTGCATGTTGGAGGGCTTCATCTCGGGTGCAGGACCCCTG 416
Db |||||
Qy 366 AGCCTTCACAGCAGTGAGTGCATGTTGGAGGGCTTCATCTCGGGTGCAGGACCCCTG 307
Db |||||
Qy 417 GGAAAGTTCAGAACTCCAGTCTCTGCTCAATTGCGCATCACTTTCAGAGCTATCA 476
Db |||||
Qy 306 GGAAAGTTCAGAACTCCAGTCTCTGCTCAATTGCGCATCACTTTCAGAGCTATCA 247
Db |||||
Qy 477 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCACAAAC 536
Db |||||
Qy 246 TGAGCCAACTCCACCCACAGGGCTCAGTCGCCACCATGTGGGCTCTCCAGTGCACAAAC 187
Db |||||
Qy 537 CACCGAGCATTCCACCATGACCGGTTCAGCTACAGTACAAATCCAGAGACCAATCTCTGCTA 596
Db |||||
Qy 186 CACCGAGCATTCCACCATGACCGGTTCAGCTACAGTACAAATCCAGAGACCAATCTCTGCTA 127
Db |||||
Qy 597 GAGTGCA 603
Db |||||
Qy 126 GAGTGCA 120
Db |||||
RESULT 13
ABL67752/c
ID ABL67752 standard; DNA; 366 BP.
XX AC ABL67752;
XX DT 15-MAY-2002 (first entry)
XX DE Oesophagus cancer related gene sequence SEQ ID NO:6089.
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX OS Homo sapiens.
XX PN WO200194629-A2.
XX PD 13-DEC-2001.
XX PF 30-MAY-2001; 2001WO-US10838.
XX PR 05-JUN-2000; 2000US-209473P.
XX PR 05-JUN-2000; 2000US-209531P.
XX PR 18-SEP-2000; 2000US-233133P.
XX PR 18-SEP-2000; 2000US-233611P.

PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
PR XX
PA (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 6089; 44pp; English.
XX
CC The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilms' tumour.
XX
SQ Sequence 366 BP; 77 A; 80 C; 119 G; 90 T; 0 other;
Query Match 24.3%; Score 196; DB 24; Length 366;

Best Local Similarity 99.6%; Pred. No. 1.5e-83;
Matches 246; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 357 AGCCTTCACAGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 416
Db |||||
Qy 366 AGCCTTCACAGCAGTGCATGTTGGAGGGCTTCATCTCGGGCTGCAAGACCCCTG 307
Db |||||
Qy 417 GGAAGTTCCAGAACTCCACGCTCCTGTCTCAATTGTGCCATCAACTTTTCAGAGCTATCA 476
Db |||||
Qy 306 GGAAGTTCCAGAACTCCACGCTCCTGTCTCAATTGTGCCATCAACTTTTCAGAGCTATCA 247
Db |||||
Qy 477 TGAGCCAACTTCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAC 536
Db |||||
Qy 246 TGAGCCAACTTCACCCACAGGGCTCAGTCGCCACCATGTGGCCCTCTCCAGTGCAAAC 187
Db |||||
Qy 537 CACGAGCATTCACCATGACCGGTACAGCTACAAATCCAGACCATCAATCTGCTA 596
Db |||||
Qy 186 CACGAGCATTCACCATGACCGGTACAGCTACAAATCCAGACCATCAATCTGCTA 127
Db |||||
Qy 597 GAGTGCA 603
Db 126 GAGTGCA 120
RESULT 14
ABQ57070/c
ID ABQ57070 standard; cDNA; 629 BP.
XX
AC ABQ57070;
XX
DT 02-AUG-2002 (first entry)
XX
DE Human colon cancer related nucleotide sequence SEQ ID NO:765.
XX
KW Human; colon cancer; cancer; tissue profiling; forensic; mapping;
KW genetic analysis; diagnostic; antisense therapy; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200229086-A2.
XX
PD 11-APR-2002.
XX
PF 02-OCT-2001; 2001WO-US30732.
XX
PR 02-OCT-2000; 2000US-237271P.
XX
PA (FARB) BAYER CORP.
XX
PI Burgess C, Astle JH, Carroll E, Catino TJ, Dwivedi P, Molino GA;
PI Thiagalingam A, Lewis ME;
XX
XX WPI; 2002-426115/45.
XX
XX New isolated nucleic acid that is differentially expressed in cancer
PT tissues useful for determining the presence of colon cancer in a cell
PT or tissue type, and in antisense therapy -
XX
PS Claim 1; Fig 1; 796pp; English.
XX
CC ABQ56306 to ABQ60787 represent isolated nucleic acids (I) differentially
CC expressed in cancer tissues. ABQ78993 to ABQ79004 represent proteins
CC encoded by the ABQ60776 to ABQ60787 nucleic acid sequences. (I) can be
CC used in antisense therapy. An antibody immunoreactive with a polypeptide
CC encoded by (I) is useful for detecting cancer in a patient sample, and
CC for detecting the presence or absence of a polynucleotide encoded by a
CC nucleic acid which hybridises to (I) in a cell. A probe/primer derived
CC from (I) can be used for determining the presence of a nucleic acid
CC hybridises to (I), and for determining the phenotype of cells in a sample
CC of cells from a patient. (I) is useful for determining the presence of
CC colon cancer in a cell or tissue type, for determining the presence or
CC state of other type of cancer, in antisense therapy, to generate
CC macroarrays on a solid surface, to identify a chromosome on which the

CC corresponding gene resides, and in tissue profiling, forensics, genetic
CC analysis, mapping and diagnostic applications. (I) can be used to raise
CC antibodies, and to screen for peptide analogues and antagonists.

XX Sequence 629 BP; 128 A; 113 C; 149 G; 198 T; 41 other;

Query Match 16.5%; Score 133; DB 24; Length 629;
Best Local Similarity 100.0%; Pred. No. 1.8e-53;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 633 AAGACTGCAGAGTCCTCCATCTTCAGTCCATTCAGCTCCCTGGCATTTAACTACCAG 692

Db 167 AAGACTGCAGAGTCCTCCATCTTCAGTCCATTCAGCTCCCTGGCATTTAACTACCAG 108

Qy 693 CATCAGTGGTCCCAAGGAATCCCTTCTAGCCTCTGACATGATCTGCTGGAAGAG 752

Db 107 CATCAGTGGTCCCAAGGAATCCCTTCTAGCCTCTGACATGATCTGCTGGAAGAG 48

Qy 753 CATCAAAACAAAC 765

Db 47 CATCAAAACAAAC 35

RESULT 15

AAD27026
ID AAD27026 standard; DNA; 4159 BP.

XX AC AAD27026;

DT 09-APR-2002 (first entry)

XX DE Human Secreted epithelial colon stromal-1 (Secs-1) gene.

XX KW Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
XX KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
XX KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
XX KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
XX KW anorectic; immunomodulator; antipsoriatic; vulnerable; anti-infectivity;
XX KW gynaecological; antiulcer; anti-inflammatory; cancer; cell therapy;
XX KW human; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 1..4159
FT /tag= a
FT /product= "Human Secs-1 protein"
FT /note= "CDS does not include stop codon"
FT /partial
FT exon 1..69
FT /tag= d
FT /note= "Encodes Secs-1 peptide (AAE16484)"
FT intron 70..2626
FT /tag= d
FT /cons_splice= (5'site:NO, 3'site:NO)
FT exon 2627..2725
FT /tag= d
FT /note= "Encodes Secs-1 peptide (AAE16485)"
FT intron 2726..4075
FT /tag= d
FT /cons_splice= (5'site:NO, 3'site:NO)
FT exon 4076..4159
FT /tag= d
FT /note= "Encodes Secs-1 peptide (AAE16486)"

XX WO200198497-A1.

XX 27-DEC-2001.

XX 28-NOV-2000; 2000WO-US32479.

XX 21-JUN-2000; 2000US-0599087.

XX 28-NOV-2000; 2000US-0724000.

XX

PA (AMGE-) AMGEN INC.

XX PI Polverino AJ, Luethy R;

XX WI; 2002-122281/16

DR P-PSDB; AAE16481, AAE16484, AAE16485, AAE16486.

XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
PT useful for diagnosing, treating and preventing hematopoietic disorder,
PT osteoporosis, Paget's disease, cancer, diabetes -
XX Disclosure; Fig 4; 134pp; English.

XX The present invention relates to an isolated murine or human secreted
CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
CC therapy and cell therapy. Secs-1 is useful for identifying a compound
CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
CC preventing or ameliorating a disease condition such as haematopoietic
CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
CC is also useful for diagnosing a pathological condition which involves
CC determining the presence or amount of Secs-1 or polypeptide encoded by
CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
CC susceptibility to pathological condition based on the presence or amount
CC of expression of the polypeptide. The present sequence is human Secs-1
CC gene.

XX SQ Sequence 4159 BP; 1009 A; 1037 C; 983 G; 1110 T; 20 other;

Query Match 11.5%; Score 93; DB 24; Length 4159;

Best Local Similarity 100.0%; Pred. No. 2.2e-14;

Matches 93; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 AGGGAAGAGGCGTCTCTGCCAAGGCTGGTCAGGAGAGAACACAGGCTCTCTGCCACCG 156

Db 2634 AGGGAAGAGGCGTCTCTGCCAAGGCTGGTCAGGAGAGAACACAGGCTCTCTGCCACCG 2693

Qy 157 AGTCCCTAGCCCCCAACTCAACAACTGAAAGG 189

Db 2694 AGTCCCTAGCCCCCAACTCAACAACTGAAAGG 2726

Search completed: May 17, 2003, 15:58:41

Job time : 276 secs

THIS PAGE BLANK (user)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 03:39:29 ; Search time 1728 Seconds
(without alignments)
759.163 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 442
Sequence: 1 MRLVLSSLLCILLCSIF.....PCKLEPRRLWVPGALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_apool/US09724000/runat_16052003_145500_17160/app_query.fasta_1.263
-DB=EST_QPWT-fasCap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09724000.CGN 1 1 1906 @runat_16052003_145500_17160 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estov:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hctc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hctc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	442	100.0	356	10	AW351839
2	418	94.6	503	9	AA422178
3	416	94.1	906	12	BE899580
4	356	80.5	557	14	BQ189412
5	350.5	79.3	399	9	AA283751
6	245	55.4	506	13	BM446173
7	245	55.4	537	12	BF041606
8	228.5	51.7	342	9	AA265120
9	228.5	51.7	436	9	AA840147
10	228.5	51.7	441	9	AA272330
11	228.5	51.7	442	9	AA242158
12	228.5	51.7	447	9	AA265314
13	228.5	51.7	450	10	BB864020
14	228.5	51.7	451	10	AW989615
15	228.5	51.7	452	9	AA222093
16	228.5	51.7	456	9	AA267641
17	228.5	51.7	456	9	AA529197
18	228.5	51.7	462	9	AA017989
19	228.5	51.7	498	9	AA240277
20	228.5	51.7	515	9	AA265055
21	228.5	51.7	515	9	AA274151
22	228.5	51.7	517	9	AA230541
23	228.5	51.7	521	9	AA272697
24	228.5	51.7	523	9	AA272358
25	228.5	51.7	543	9	AA538434
26	228.5	51.7	549	9	AA220337
27	228.5	51.7	554	9	AA538362
28	228.5	51.7	563	9	AA238890
29	228.5	51.7	753	13	BG966710
30	228.5	51.7	758	11	AK012157
31	228.5	51.7	800	14	BQ931917
32	225.5	51.0	456	9	AA498458
33	225.5	51.0	907	12	BF531305
34	220.5	49.9	742	13	BF580962
35	220.5	49.9	765	12	BF580962
36	212.5	48.1	444	9	AA459817
37	205.5	46.5	267	9	AA509491
38	205.5	46.5	304	14	W08316
39	199.5	45.1	529	9	AA463145
40	199	45.0	422	10	AW142726
41	195	44.1	492	10	AW913863
42	195	44.1	510	14	BQ703407
43	195	44.1	531	12	BG374322
44	194.5	44.0	466	9	AA528993
45	182.5	41.3	1044	12	BF582268

ALIGNMENTS

RESULT 1
LOCUS AW351839 356 bp mRNA linear EST 01-FEB-2000
DEFINITION RC1-CT0199-150999-021-A03 CT0199 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW351839
VERSION AW351839.1 GI:6849552
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 356)
AUTHORS HCGP <http://www.ludwig.org.br/ORESTES>.
TITLE The FAPESP/LICR Human Cancer Genome Project
JOURNAL Unpublished (1999)


```

VERSION      BE899580.1  GI:10367234
KEYWORDS     EST.
SOURCE       human.
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 906)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE        Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL      NIH-MGC http://mgc.nci.nih.gov/.
COMMENT      National Institutes of Health, Mammalian Gene Collection (MGC)
              Unpublished (1999)
              Contact: Robert Strausberg, Ph.D.
              Email: cgapbs@mail.nih.gov
              Tissue Procurement: DCTD/DTF
              cDNA Library Preparation: Ling Hong/Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
              Plate: LUCM822 row: f column: 12
              High quality sequence start: 23
              High quality sequence stop: 775.
              Location/Qualifiers
FEATURES     source
             1..906
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone_image="3952523"
             /clone_lib="NIH MGC 9"
             /tissue_type="adenocarcinoma cell line"
             /lab_host="DH10B (phage-resistant)"
             /note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:
             EcoRI; cDNA made by oligo-dt priming. Directionally
             cloned into EcoRI/XhoI sites using the following 5'
             adaptor: GGCACGAG(G). Size-selected >500bp for average
             insert size 1.8kb. Library constructed by Ling Hong in
             the laboratory of Gerald M. Rubin (University of
             California, Berkeley) using ZAP-cDNA synthesis kit
             (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT   219 a 282 c 219 g 186 t
ORIGIN
Alignment Scores:
Pred. No.:    1.78e-33      Length:      906
Score:        416.00      Matches:    81
Percent Similarity: 97.59%      Conservative: 0
Best Local Similarity: 97.59%      Mismatches: 0
Query Match:  94.12%      Indels:      2
DB:           12          Gaps:        0

US-09-724-000A-5 (1-81) x BE899580 (1-906)
Qy  1 MetArgLeuLeu-VallLeuSerSerLeuLeu-CysIleLeuLeuLeuCysPheSerIleP 20
Db  69 ATGAGGCTTCTAGGTCCTTTCAGCGCTGCTTGTATCTCTGCTTCTGCTTCTCCATCT 128
Qy  20 heSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuC 40
Db  129 TCTCCACAGAAGGGAAGAGGGCTCTCTCAAGGCGCTGTCAGCGAGGAGAACCCAGGCTCT 188
Qy  40 yScyHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuC 60
Db  189 GCTGCCACCGAGTCCCTTAGCCCCCACTCAACAAACCTGAAAGGACATCATGTGAGGCTCT 248
Qy  60 yLysProCysLysLeuGluProGluProArgLeuTrpValValProGlyAlaLeuProG 80
Db  249 GTAACCATGCAAGCTTGAGCGACAGAGCCCCCGCTTTGGGTGGTGGCGACTCCAC 308
Qy  80 InVal 81
Db  309 AGGTG 313

RESULT 4
BQ189412
LOCUS      BQ189412          557 bp      mRNA      linear      EST 30-APR-2002

DEFINITION UI-E-EJ1-ajv-h-16-0-UI.r1 UI-E-EJ1 Homo sapiens cDNA clone
ACCESSION  BQ189412
VERSION    BQ189412.1  GI:20364963
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 557)
AUTHORS    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL    Bonaldo M.F., Lennon G. and Soares M.B.
COMMENT    Normalization and subtraction: two approaches to facilitate gene
              discovery
              Genome Res. 6 (9), 791-806 (1996)
              9704477
              Contact: Soares, MB
              Program for Rat Gene Discovery and Mapping
              University of Iowa
              451 Eckstein Medical Research Building Iowa City, IA 52242, USA
              Tel: 319 335 8250
              Fax: 319 335 9565
              Email: msoares@blue.weeg.uiowa.edu
              Tissue Procurement: Dr. Gregg Hageman
              cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
              cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
              DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
              Clone Distribution: Researchers may obtain clones from Research
              Genetics (www.resgen.com).
              Seq primer: M13 REVERSE.
              Location/Qualifiers
FEATURES     source
             1..557
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="UI-E-EJ1-ajv-h-16-0-UI"
             /clone_lib="UI-E-EJ1"
             /tissue_type="fetal eyes, lens, eye anterior segment,
             optic nerve, retina, Retina foveal and Macular, RPE and
             Choroid"
             /dev_stage="fetal and adult"
             /lab_host="DH10B (Life Technologies) (Ti phage resistant)"
             /note="Organ: eye; Vector: p773-Pac (Pharmacia) with a
             modified polylinker; Site 1: EcoR I; Site 2: Not I;
             UI-E-EJ1 is a substracted cDNA library constructed
             according to Bonaldo, Lennon and Soares, Genome Research,
             6:791-806, 1996. First strand cDNA synthesis was primed
             with an oligo-dT primer containing a Not I site. Double
             stranded cDNA was ligated to an EcoR I adaptor, digested
             with Not I, and cloned directionally into p773-Pac
             vector. The oligonucleotide used to prime the synthesis of
             first-strand cDNA contains a library tag sequence that is
             located between the Not I site and the (dT)18 tail. The
             sequence tags for this library are: fetal eyes, AGAATCAAGA
             ; lens, CGATTAGCGA; eye anterior segment, AATGCCGAT;
             optic nerve, CCATTAGTG; retina, CCGG; Retina foveal and
             Macular, GTCC; RPE and Choroid, ACCTA. This library was
             created for the program, Gene Discovery in the Visual
             System, supported by National Eye Institute (NEI)."
BASE COUNT   135 a 182 c 138 g 101 t
ORIGIN
Alignment Scores:
Pred. No.:    1.78e-27      Length:      557
Score:        356.00      Matches:    63
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:  80.54%      Indels:      0
DB:           14          Gaps:        0

US-09-724-000A-5 (1-81) x BQ189412 (1-557)
Qy  19 IlePheSerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgThrArg 38
Db  1 ATCTTCTCCAGAGGGAAGAGGGCTCTCTCAAGGCGCTGTCAGCGAGGAGAACCCAGG 60

```


US-09-724-000A-5 (1-81) x BM446173 (1-506)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysileLeuLeuLeuLeuCysPheSerilePhe 20
 Db 39 ATGAGGCTTCTAATCCTCACCAGCCTGCTGTCATCTCTCTGCTTCTGCTTCTCGCTCTTC 98
 Qy 21 SerThrGluGlyLys-----ArgArgProAlaLysAlaTrpSerGlyArgArgThr 37
 Db 99 TCCGAGAAGGAGAGTCACTGAGGATCATGCCAAGCCCGGAAAGGCAAGCC--- 155
 Qy 38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisVal 57
 Db 156 -----TGCTGTCCCGAATTCCTGGCCCTGACCTGATGACCCAGAAAGGACCGTACG 209
 Qy 58 ArgLeuCysLysProCysGlyLeuGluProGluProArgLeuTrpValProGlyVala 77
 Db 210 AGAAACTGCAGACCATGCAAACTCAAGTCAAAGCACCCTTTGGTGGTTCCTGGGGCA 269
 Qy 78 LeuProGlnVal 81
 Db 270 CTCCACAGGTG 281

RESULT 7
 BF041606
 LOCUS
 DEFINITION BP250007A20H3 Soares normalized bovine placenta Bos taurus cDNA
 clone BP250007A20H3 5', mRNA sequence.

ACCESSION
 VERSION BF041606
 KEYWORDS BF041606.1 GI:10758652
 SOURCE EST.

ORGANISM

Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.

REFERENCE
 AUTHORS Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson
 J.H.

TITLE

Bovine ESTs

Unpublished (2000)

Contact: Lewin, H. A.

W. M. Keck Center for Comparative and Functional Genomics

University of Illinois at Urbana-Champaign

340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL

61801, USA

Tel: 217 333 5998

Fax: 217 244 5617

Email: h-lewin@uiuc.edu

Funding for cattle EST sequencing was provided by the USDA National
 Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
 to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
 from Washington University Genome Center. Vector Trimmi g.

Cross match from Washington University Genome Center PHRAP suite.

This sequence is vector free and at least 200 bp in length.

PCR Primers

FORWARD: TAATACGACTCACTATAGGG

BACKWARD: ATTAACCTCACTAAAG

Insert Length: 537 Std Error: 0.00

Plate: BP250007A20 row: H column: 3

Seq primer: AGCGATAACAATTCACAGGA

High quality sequence stop: 537.

Location/Qualifiers

1. .537

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone="BP250007A20H3"

/clone_lib="Soares normalized bovine placenta"

/sex="Female"

/lab_host="DH10B"

/note="Organ; placenta; Vector: pT73pac; Site 1: EcoRI;
 Site 2: NotI; The cDNA library was contributed by the

Soares laboratory and it was constructed and normalized
 as described by Bonaldo, M.F., Lennon, G. and Soares,

BASE COUNT 128 a 166 c 120 g 122 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 7,81e-16 Length: 537
 Score: 245.00 Matches: 48
 Percent Similarity: 69.05% Conservative: 10
 Best Local Similarity: 57.14% Mismatches: 20
 Query Match: 55.43% Indels: 6
 DB: 12 Gaps: 2

US-09-724-000A-5 (1-81) x BF041606 (1-537)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysileLeuLeuLeuLeuCysPheSerilePhe 20
 Db 28 ATGAGGCTTCTAATCCTCACCAGCCTGCTGTCATCTCTGCTTCTCTGCTTCTCGCTTC 87

Qy 21 SerThrGluGlyLys-----ArgArgProAlaLysAlaTrpSerGlyArgArgThr 37
 Db 88 TCCGAGAAGGAGAGTCACTGAGGATCATGCCAAGCCCGGAAAGGCAAGCC--- 144

Qy 38 ArgLeuCysCysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisVal 57
 Db 145 -----TGCTGTCCCGAATTCCTGGCCCTGACCTGATGACCCAGAAAGGACCGTACG 198

Qy 58 ArgLeuCysLysProCysGlyLeuGluProGluProArgLeuTrpValProGlyVala 77
 Db 199 AGAACTGCAGACCATGCAAACTCAAGTCAAAGCACCCTTTGGTGGTTCCTGGGGCA 258

Qy 78 LeuProGlnVal 81

Db 259 CTCCACAGGTG 270

RESULT 8

AA265120

LOCUS

DEFINITION

AA265120

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

REFERENCE

AUTHORS

SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 442)
REFERENCE Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine#
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:430909
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 388.
FEATURES
Location/Qualifiers
1..442
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:697349"
 /clone_lib="Barstead mouse pooled organs MPLRB4"
 /sex="mixed"
 /tissue.type="pooled organs"
 /dev_stage="7 day"
 /lab_host="DH10B"
 /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
 a modified polylinker; Site 1: EcoRI; Site 2: NotI; 1st
 strand cDNA was primed with a Not I - oligo(dt) primer [5'
 TGTTACGACTGAAGTCGGAGCGCCGCCCTTTTTTTTTTTTTTTTTT
 3']; double-stranded cDNA was ligated to Eco RI adaptors
 [GTTCAGATTCGGTACC], digested with Not I and cloned into
 the Not I and Eco RI sites of the modified pT7T3 vector.
 Library constructed by Bob Barstead."
BASE COUNT 103 a 134 c 102 g 103 t
ORIGIN

Alignment Scores:
Pred. No.: 3,2e-14 Length: 442
Score: 228.50 Matches: 45
Percent Similarity: 67.90% Conservative: 10
Best Local Similarity: 55.56% Mismatches: 23
Query Match: 51.70% Indels: 3
DB: 9 Gaps: 1

US-09-724-000A-5 (1-81) x AA242158 (1-442)

QY 1 MetArgLeuValLeuSerSerLeuLeuCysAlaLeuLeuLeuLeuLeuLeuPhe 20
Db 52 ATGAGACTTCTAGCCCTTTCCGGCTGCCTCTGCATGCTGCTCTCTGCAATTTTC 111
QY 21 SerThrGlucLyLysArgArgProAlaValAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 112 TCTCAGAGGGAGAGACATCTTGCAAGTCTCTGAACTCAGGC-----TGC 162
QY 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisIshisValArgLeuCys 60
Db 163 TGTCACCTATCTCTTAGATCCAAGCTGCACAACCTGGAAAGAAAACACACAGCCCTGC 222
QY 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 223 AGACTTCGCAAAACAAGCTTACCAGTCAAGTCATGGGTGGTCTGGGGCTCTCCCACAG 282
OY 81 Val 81

cdna"
/tissue_type="kidney"
/cell_line="CCL-142 RAG"
/note="pooled cell lines ; (cell_line=RCB-1751 WEHI 164),
(cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),
(cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
(cell_line=RCB-0559 K-1 . Fl), (cell_line=RCB-1283 B16
melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
(cell_type=Leidyig cells, cell_line=CRL-2065 MLTC-1),
(cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
(tissue_type=bladder, cell_line=RCB-0544 MBT-2),
(tissue_type=bone marrow, cell_type=stroma cell,
cell_line= CRL-2028 SR-4987), (tissue_type=colon,
cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C,
cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
(, (strain=C3H, tissue_type=brain, cell_line=CRL-1443
BC3H1)"

BASE COUNT 101 a 136 c 107 g 106 t
ORIGIN

Alignment Scores:
Pred. No.: 3.28e-14 Length: 450
Score: 228.50 Matches: 45
Percent Similarity: 67.90% Conservative: 10
Best Local Similarity: 55.56% Mismatches: 23
Query Match: 51.70% Indels: 3
DB: 10 Gaps: 1

US-09-724-000A-5 (1-81) x BB864020 (1-450)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db 57 ATGAGACTTCCTAGCCCTTTCCGGTCTGCTCGCAGCTGCTCTCTGTTTCTGCATTTTC 116
Qy 21 SerThrGluGlyArgArgProAlaLysAlaTrpSerGlyArgArgThArgLeuCys 40
Db 117 TCCTCAGAAGGGAGAAGACATCTCTCCCAAGTCTTGAAGAACTCAGCGCG------TGC 167
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys 60
Db 168 TGTCACTTATCTCTAGATCCCAAGCTGACACCTCGGAAGGAAGAAACACACAAAGGCCCTGC 227
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 228 AGACTCTCCAGAACAACTACCAGTCAAGTCATGGGTGGTGGCTGGGGCTCTCCACAG 287
Qy 81 Val 81
Db 288 ATA 290

RESULT 14
AW989615
LOCUS uf18c02.y1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
DEFINITION IMAGE:1511714 5', mRNA sequence.
ACCESSION AW989615
VERSION AW989615.1 GI:8185064
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 451)
REFERENCE NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
TITLE Unpublished (1997)
JOURNAL
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:938566
Seq primer: -40RP from Gibco
High quality sequence stop: 451.
FEATURES
source
1..451
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:1511714"
/sex="Female (lactating)"
/tissue_type="mammary gland"
/lab_host="DH10B"
/note="vector: p773D-Pac (Pharmacia) with a modified
polylinker; 1st strand cDNA was prepared from mammary
gland tissue from a lactating female, and was then primed
with a Not I - oligo(dT) primer. Double-stranded cDNA was
ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the
modified p773 vector. Library is normalized. Library
was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 105 a 137 c 102 g 107 t
ORIGIN

Alignment Scores:
Pred. No.: 3.29e-14 Length: 451
Score: 228.50 Matches: 45
Percent Similarity: 67.90% Conservative: 10
Best Local Similarity: 55.56% Mismatches: 23
Query Match: 51.70% Indels: 3
DB: 10 Gaps: 1

US-09-724-000A-5 (1-81) x AW989615 (1-451)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePhe 20
Db 38 ATGAGACTTCCTAGCCCTTTCCGGTCTGCTCGCAGTCTCTCTGTTTCTGCATTTTC 97
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThArgLeuCys 40
Db 98 TCCTCAGAAGGGAGAAGACATCTCTCCCAAGTCTTGAAGAACTCAGCGCG------TGC 148
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys 60
Db 149 TGTCACTTATCTCTAGATCCCAAGCTGACACCTCGGAAGGAAGAAACACACAAAGGCCCTGC 208
Qy 61 LysProCysLysLeuGluProGluProArgLeuTrpValProGlyAlaLeuProGln 80
Db 209 AGACTCTCCAGAACAACTACCAGTCAAGTCATGGGTGGTGGCTGGGGCTCTCCACAG 268
Qy 81 Val 81
Db 269 ATA 271

RESULT 15
AA222093
LOCUS my29b11.r1 Barstead mouse pooled organs MRLRB4 Mus musculus cDNA
DEFINITION clone IMAGE:697245 5', mRNA sequence.
ACCESSION AA222093
VERSION AA222093.1 GI:1841618
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 452)
REFERENCE Garra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:35:10 ; Search time 20 Seconds
(without alignments)
390.626 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLVLSSLICILLCSIF.....PCKLEPEPRLWVFGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 362588 seqs, 96450795 residues

Word size : 0

Total number of hits satisfying chosen parameters: 47966

Minimum DB seq length: 23

Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6	7.4	28	9 US-09-749-637A-265	Sequence 265, App
2	6	7.4	54	10 US-09-864-761-47036	Sequence 47036, A
3	6	7.4	65	9 US-10-091-504-786	Sequence 786, App
4	6	7.4	65	10 US-09-764-869-786	Sequence 786, App
5	6	7.4	72	10 US-09-864-761-36193	Sequence 36193, A
6	6	7.4	78	10 US-09-864-761-37299	Sequence 37299, A
7	6	7.4	79	9 US-10-083-357-1164	Sequence 1164, App
8	5	6.2	23	8 US-08-424-5508-400	Sequence 400, App
9	5	6.2	24	9 US-09-843-221A-132	Sequence 132, App
10	5	6.2	24	9 US-09-843-221A-142	Sequence 142, App
11	5	6.2	24	9 US-10-150-111-127	Sequence 127, App
12	5	6.2	24	10 US-09-242-980-15	Sequence 15, App
13	5	6.2	24	12 US-10-001-879-157	Sequence 157, App
14	5	6.2	25	9 US-09-989-920-214	Sequence 214, App
15	5	6.2	26	10 US-09-925-300-1735	Sequence 1735, App
16	5	6.2	27	9 US-10-083-357-1099	Sequence 1099, App
17	5	6.2	27	10 US-09-864-761-39998	Sequence 39998, A
18	5	6.2	27	10 US-09-864-761-40419	Sequence 40419, A
19	5	6.2	27	10 US-09-911-888-8	Sequence 8, Appli

20	5	6.2	28	9 US-09-843-221A-96	Sequence 96, Appl
21	5	6.2	28	9 US-09-843-221A-106	Sequence 106, App
22	5	6.2	28	10 US-09-864-761-41850	Sequence 41850, A
23	5	6.2	29	9 US-09-362-179-5	Sequence 5, Appli
24	5	6.2	29	10 US-09-864-761-44425	Sequence 44425, A
25	5	6.2	30	9 US-09-843-221A-127	Sequence 127, App
26	5	6.2	30	9 US-09-843-221A-137	Sequence 137, App
27	5	6.2	30	9 US-09-843-221A-156	Sequence 156, App
28	5	6.2	30	12 US-10-127-318-5	Sequence 5, Appli
29	5	6.2	31	10 US-09-142-755-10	Sequence 10, Appli
30	5	6.2	32	9 US-09-983-802-207	Sequence 207, App
31	5	6.2	32	10 US-09-922-261-157	Sequence 157, App
32	5	6.2	32	10 US-09-949-695A-20	Sequence 20, Appl
33	5	6.2	33	10 US-09-975-901-17	Sequence 17, Appl
34	5	6.2	34	9 US-09-843-221A-91	Sequence 91, Appl
35	5	6.2	34	9 US-09-843-221A-101	Sequence 101, App
36	5	6.2	34	9 US-09-843-221A-120	Sequence 120, App
37	5	6.2	35	10 US-09-864-761-44044	Sequence 44044, A
38	5	6.2	35	10 US-09-922-261-155	Sequence 155, App
39	5	6.2	36	9 US-09-764-891-2801	Sequence 2801, App
40	5	6.2	36	10 US-09-864-761-35960	Sequence 35960, A
41	5	6.2	36	10 US-09-864-761-36141	Sequence 36141, A
42	5	6.2	36	10 US-09-867-852-20	Sequence 20, Appl
43	5	6.2	37	10 US-09-242-980-14	Sequence 14, Appl
44	5	6.2	38	9 US-09-984-345-207	Sequence 207, App
45	5	6.2	38	9 US-09-966-362-207	Sequence 207, App

ALIGNMENTS

RESULT 1
US-09-749-637A-265
; Sequence 265, Application US/09749637A
; Patent No. US20020173449A1
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: Cartier, G. Edward
; APPLICANT: Watkins, Maren
; APPLICANT: Hilliard, David R.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Layer, Richard T.
; APPLICANT: Jones, Robert M.
; TITLE OF INVENTION: O-Superfamily Conotoxin Peptides
; FILE REFERENCE: 2314-227
; CURRENT APPLICATION NUMBER: US/09749,637A
; CURRENT FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/243,412
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US60/219,440
; PRIOR FILING DATE: 2000-07-20
; PRIOR APPLICATION NUMBER: US 60/214,263
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/173,754
; PRIOR FILING DATE: 1999-12-30
; NUMBER OF SEQ ID NOS: 409
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 265
; LENGTH: 28
; TYPE: PRT
; ORGANISM: Conus delessertii
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1)...(28)
; OTHER INFORMATION: Xaa at residues 4 and 14 may be Pro or hydroxy-Pro; Xaa at residues 15 and 21 may be Glu or gamma-carboxy-Glu; Xaa at residues 15 and 21 may be Tyr, 125-I-Tyr, mono-iodo-Tyr, di-iodo-Tyr, O-sulpho-Tyr or O-phospho-Tyr
US-09-749-637A-265

Query Match 7.4%; Score 6; DB 9; Length 28;

Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 CILLIC 16
| | | | |
Db 22 CILLIC 27

RESULT 2
US-09-864-761-47036
; Sequence 47036, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47036
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005317.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.68
; OTHER INFORMATION: EST_HUMAN HIT: AW500861.1, EVALUATE 2.00e-15
US-09-864-761-47036

Query Match 7.4%; Score 6; DB 10; Length 54;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LVLSS 8
| | | | |
Db 35 LVLSS 40

RESULT 3
US-10-091-504-786
; Sequence 786, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 786
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-091-504-786

Query Match 7.4%; Score 6; DB 9; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LVLSS 8
| | | | |
Db 6 LVLSS 11

RESULT 4
US-09-764-869-786
; Sequence 786, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 786
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: SITE
; LOCATION: (17)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-786

Query Match 7.4%; Score 6; DB 10; Length 65;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LVLSS 8
 |||||
Db 6 LVLSS 11

RESULT 5
US-09-864-761-36193
; Sequence 36193, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36193
LENGTH: 72
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
; OTHER INFORMATION: MAP TO AC012127.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: O61608, EVALUE 7.60e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE789526.1, EVALUE 8.10e-01
US-09-864-761-36193
Query Match 7.4%; Score 6; DB 10; Length 72;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 10 LCILL 15
 |||||
Db 1 LCILL 6
RESULT 6
US-09-864-761-37299
; Sequence 37299, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 37299
LENGTH: 78
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
; OTHER INFORMATION: MAP TO AC010906.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
US-09-864-761-37299

Query Match 7.4%; Score 6; DB 10; Length 78;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LVLSS 8
Db 71 LVLSS 76

RESULT 7
US-10-083-357-1164
; Sequence 1164, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiangong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083.357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 1164
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-1164

Query Match 7.4%; Score 6; DB 9; Length 79;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLVLS 7
Db 67 RLVLS 72

RESULT 8
US-08-424-550B-400
; Sequence 400, Application US/08424550B
; Patent No. US20020119447A1
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D177/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMBSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-424-550B-400

Query Match 6.2%; Score 5; DB 8; Length 23;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GKRRP 28
Db 8 GKRRP 12

RESULT 9
US-09-843-221A-132
; Sequence 132, Application US/09843221A
; Publication No. US20030039654A1
; GENERAL INFORMATION:
; APPLICANT: KOSTENUK, PAUL
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: LACEY, DAVID LEE
; TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID
; TITLE OF INVENTION: RELATED PROTEIN
; FILE REFERENCE: A-665B
; CURRENT APPLICATION NUMBER: US/09/843,221A
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/266,673
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/214,860
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/200,053
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 132
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified human PTH
US-09-843-221A-132

Query Match 6.2%; Score 5; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLL 10
Db 18 LSSLL 22

RESULT 10
US-09-843-221A-142
; Sequence 142, Application US/09843221A
; Publication No. US20030039654A1


```

; GENERAL INFORMATION:
; APPLICANT: KOSTENUK, PAUL
; APPLICANT: LIU, CHUAN-FA
; APPLICANT: LACEY, DAVID LEE
; TITLE OF INVENTION: MODULATORS OF RECEPTORS FOR PARATHYROID HORMONE AND PARATHYROID H
; TITLE OF INVENTION: RELATED PROTEIN
; FILE REFERENCE: A-665B
; CURRENT APPLICATION NUMBER: US/09/843,221A
; CURRENT FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 60/266,673
; PRIOR FILING DATE: 2001-02-06
; PRIOR APPLICATION NUMBER: 60/214,860
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/200,053
; PRIOR FILING DATE: 2000-04-27
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 142
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified human PTHrP
US-09-843-221A-142

Query Match          6.2%; Score 5; DB 9; Length 24;
Best Local Similarity 100.0%; Pred.No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLL 10
Db 18 LSSLL 22

RESULT 11
US-10-150-111-127
; Sequence 127, Application US/10150111
; Publication No. US20030078386A1
; GENERAL INFORMATION:
; APPLICANT: Rubin et al.
; TITLE OF INVENTION: Secreted Protein HPEAD48
; FILE REFERENCE: P2018P1D1
; CURRENT APPLICATION NUMBER: US/10/150,111
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/288,143
; PRIOR FILING DATE: 1999-04-08
; PRIOR APPLICATION NUMBER: PCT/US98/21142
; PRIOR FILING DATE: 1998-10-08
; PRIOR APPLICATION NUMBER: 60/061,463
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,529
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/071,498
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,527
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,536
; PRIOR FILING DATE: 1997-10-09
; PRIOR APPLICATION NUMBER: 60/061,532
; PRIOR FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 127
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-150-111-127

Query Match          6.2%; Score 5; DB 9; Length 24;
Best Local Similarity 100.0%; Pred.No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSS 8
Db 12 LVLSS 16
```

```

Db 9 LVLSS 13

RESULT 12
US-09-242-980-15
; Sequence 15, Application US/09242980
; Patent No. US20010021073A1
; GENERAL INFORMATION:
; APPLICANT: Ogi, Kasuhiro
; APPLICANT: Onda, Haruo
; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE
; FILE REFERENCE: 2417USOP
; CURRENT APPLICATION NUMBER: US/09/242,980
; CURRENT FILING DATE: 1999-02-28
; PRIOR APPLICATION NUMBER: PCT/JP97/03194
; PRIOR FILING DATE: 1997-08-10
; PRIOR APPLICATION NUMBER: JP 8-240880
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: JP 8-318049
; PRIOR FILING DATE: 1996-11-28
; PRIOR APPLICATION NUMBER: JP 9-135633
; PRIOR FILING DATE: 1997-05-27
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE:
; SEQ ID NO 15
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Human
US-09-242-980-15

Query Match          6.2%; Score 5; DB 10; Length 24;
Best Local Similarity 100.0%; Pred.No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 VPGAL 78
Db 10 VPGAL 14

RESULT 13
US-10-001-879-157
; Sequence 157, Application US/10001879
; Patent No. US20020127237A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Ali, Shujath
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Prostate Specific Genes and
; FILE REFERENCE: DEX-0281
; CURRENT APPLICATION NUMBER: US/10/001,879
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,188
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 201
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 24
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-879-157

Query Match          6.2%; Score 5; DB 12; Length 24;
Best Local Similarity 100.0%; Pred.No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSS 8
Db 12 LVLSS 16
```

RESULT 14

US-09-989-920-214
; Sequence 214, Application US/09989920
; Patent NO. US20020172957A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Chen, Sei-yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Proteins
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 214
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-989-920-214

Query Match 6.2%; Score 5; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 6 LSSLL 10
| | | | |
Db 20 LSSLL 24

RESULT 15

US-09-925-300-1735
; Sequence 1735, Application US/09925300
; Patent NO. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1735
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-300-1735

Query Match 6.2%; Score 5; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 43 RVPSP 47
| | | | |
Db 4 RVPSP 8

Search completed: May 16, 2003, 15:36:37
Job time : 29 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:34:10 ; Search time 16 Seconds
(without alignments)
486.681 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLVLSSLLCILLCSIF.....PCKLEPPRLWVPGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 21742

Minimum DB seq length: 23

Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : PIR 73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	7.4	52	2 G84477	hypothetical prote
2	6	7.4	77	2 T49620	hypothetical prote
3	6	7.4	79	2 A82667	single-stranded DN
4	5	6.2	23	2 D33174	trimatin - wheat (
5	5	6.2	26	2 D49533	T-cell receptor be
6	5	6.2	26	2 C33174	avematin - oat (fr
7	5	6.2	26	4 I55277	hemoglobin alpha c
8	5	6.2	28	2 PH0231	T-cell receptor Vb
9	5	6.2	29	2 E47719	house-dust-mite-re
10	5	6.2	29	2 S58390	T-cell receptor be
11	5	6.2	31	2 PH0236	T-cell receptor Vb
12	5	6.2	33	2 I52083	major acute phase
13	5	6.2	35	2 P50128	H-2 class I histoc
14	5	6.2	41	2 B81140	50S ribosomal prot
15	5	6.2	41	2 D82269	probable ribosomal
16	5	6.2	41	2 E82556	50S ribosomal prot
17	5	6.2	42	2 D30603	T-cell receptor be
18	5	6.2	44	2 S17573	thauartin-like pro
19	5	6.2	47	2 AG0380	probable ribosomal
20	5	6.2	47	2 S70809	pilC-like protein
21	5	6.2	49	2 S13077	hypothetical prote
22	5	6.2	51	2 G84217	hypothetical prote
23	5	6.2	54	2 T42306	hypothetical prote
24	5	6.2	59	2 S01877	NADH2 dehydrogenas
25	5	6.2	59	4 T07378	probable hydroxyme
26	5	6.2	61	2 C30603	T-cell receptor be
27	5	6.2	61	2 AB1587	transcription regu
28	5	6.2	61	2 C69260	hypothetical prote
29	5	6.2	61	2 S21942	surface lectin, 17

30	5	6.2	61	2 AI3331	chiM protein [impo
31	5	6.2	62	2 S30239	hypothetical prote
32	5	6.2	62	2 H81337	probable periplasm
33	5	6.2	63	2 D82547	hypothetical prote
34	5	6.2	63	2 H89761	hypothetical prote
35	5	6.2	63	2 AC2553	hypothetical prote
36	5	6.2	64	2 E81078	hypothetical prote
37	5	6.2	64	2 S01103	hypothetical prote
38	5	6.2	66	2 T12638	homeotic protein H
39	5	6.2	66	2 D72380	hypothetical prote
40	5	6.2	68	2 D27578	T-cell receptor be
41	5	6.2	68	2 S52778	T-cell receptor be
42	5	6.2	68	2 A88950	protein K09B5.10 [
43	5	6.2	68	2 H85826	hypothetical prote
44	5	6.2	68	2 AG3335	hypothetical prote
45	5	6.2	69	2 G95002	hypothetical prote

ALIGNMENTS

RESULT 1

G84477
hypothetical protein At2g06480 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: G84477
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, I.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84477
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-52 <STO>
A;Cross-references: GB:AE002093; NID:g4594353; PIDN:AAD25147.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g06480
A;Map position: 2

Query Match 7.4%; Score 6; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 VPGALP 79
Db 2 VPGALP 7

RESULT 2

T49620
hypothetical protein B5022.20 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49620
R;Schulte, U.; Algn, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49620
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-77 <SCH>
A;Cross-references: EMBL:AJ355932; GSPDB:GN00116; NCSP:B5022.20
A;Experimental source: BAC clone B5022; strain OR74A
C;Genetics:
A;Gene: NCSP:B5022.20
A;Map position: 6
A;Introns: 35/2

Query Match 7.4%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 77;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LCILL 15

Db 62 LCILL 67

RESULT 3

A82667 single-stranded DNA binding protein XF158 [imported] - Xylella fastidiosa (strain 9a5c)

C/Species: Xylella fastidiosa

C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 17-Nov-2000

C/Accession: A82667

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A/Reference number: A82515; MUID:20365717; PMID:10910347

A/Note: for a complete list of authors see reference number A59328 below

A/Accession: A82667

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-79 <SIM>

A/Cross-references: GB:A8003985; GB:A8003849; NID:G9106593; PIDN:AAP84367.1; GSPDB:GN001

A/Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briões, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorzy, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

A.F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A/Reference number: A59328

A/Contents: annotation

C/Genetics:

A/Gene: XF1558

C/Superfamily: bacterial single-stranded DNA-binding protein; single-stranded DNA-bindin

Query Match 7.4%; Score 6; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 78;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 26 RRPKA 31

Db 53 RRPKA 58

RESULT 4

D33174

trimatin - wheat (fragment)

C/Species: Triticum aestivum (common wheat)

C/Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 12-Apr-1995

C/Accession: D33174

R:Vigers, A.J.; Roberts, W.K.; Selitrennikoff, C.P.

submitted to the Protein Sequence Database, May 1991

A/Reference number: A33174

A/Accession: D33174

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-23 <VIG>

C/Superfamily: thaumatin I

Query Match 6.2%; Score 5; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 PGALP 79

Db 15 PGALP 19

RESULT 5

D49533

T-cell receptor beta chain (V-D-J region, clone HD8, LCMV-specific) - mouse (fragment)

N/Alternate names: lymphocytic choriomeningitis virus-specific T-cell receptor beta cha

C/Species: Mus musculus (house mouse)

C/Date: 06-Oct-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997

C/Accession: D49533

R:Horwitz, M.S.; Yanagi, Y.; Oldstone, M.B.

J. Virol. 68, 352-357, 1994

A/Title: T-cell receptors from virus-specific cytotoxic T lymphocytes recognizing a sin

A/Reference number: A49533; MUID:94076427; PMID:7504738

A/Accession: D49533

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-26 <HOR>

A/Experimental source: BALB/c, cytotoxic T lymphocytes

A/Note: sequence extracted from NCBI backbone (NCBIN:142930, NCBIPI:142931)

C/Superfamily: immunoglobulin V region; immunoglobulin homology

C/Keywords: T-cell receptor

Query Match 6.2%; Score 5; DB 2; Length 26;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLLVL 6

Db 19 RLLVL 23

RESULT 6

C33174

avenatin - oat (fragment)

C/Species: Avena sativa (oat)

C/Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 12-Apr-1995

C/Accession: C33174

R:Vigers, A.J.; Roberts, W.K.; Selitrennikoff, C.P.

submitted to the Protein Sequence Database, May 1991

A/Reference number: A33174

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-26 <VIG>

C/Superfamily: thaumatin I

Query Match 6.2%; Score 5; DB 2; Length 26;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 75 PGALP 79

Db 15 PGALP 19

RESULT 7

I55277

hemoglobin alpha chain thalassemia mutant Brooklyn - human (fragment)

C/Species: Homo sapiens (man)

C/Date: 24-May-1996 #sequence_revision 31-May-1996 #text_change 20-Apr-2000

C/Accession: I55277

R:Safaya, S.; Rieder, R.F.

J. Biol. Chem. 263, 4328-4332, 1988

A/Title: Dysfunctional alpha-globin gene in hemoglobin H disease in blacks. A dinucleot

A/Reference number: I55277; MUID:98153756; PMID:2831226

A/Accession: I55277

A/Molecule type: DNA

A/Residues: 1-26 <SAF>

A/Cross-references: GB:M22814; NID:G183799; PIDN:AAA52631.1; PID:G183801

A/Note: GenBank entry HUMHBADY2 repeats the normal sequence, which was not redetermined

C/Genetics:

A/Gene: GDB:HBA2

A/Cross-references: GDB:119294; OMIM:141850

A;Map position: 16p13.3-16p13.3

Query Match 6.2%; Score 5; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 67 PEPRL 71
|||||
Db 19 PEPRL 23

RESULT 8

PH0231

T-cell receptor Vb CDR3, carrier PBL Vb 6.sbt - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 30-May-1997
C;Accession: PH0231
R;Hara, H.; Morita, M.; Iwaki, T.; Hatae, Y.; Kitamoto, T.; Akizuki, S.; Goto, submitted to JIPID, June 1994
A;Description: Detection of HTLV-I proviral DNA and analysis of T cell receptor Vb CDR3
A;Reference number: PH0227
A;Accession: PH0231
A;Molecule type: mRNA
A;Residues: 1-28 <HAR>
A;Experimental source: spinal cord
C;Genetics:
A;Map position: 7
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: receptor

Query Match 6.2%; Score 5; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLLVL 6
|||||
Db 19 RLLVL 23

RESULT 9

E47719

house-dust-mite-reactive T-cell receptor beta chain (CD4+ clone DE26, V(D)J junctional
C;Species: Homo sapiens (man)
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 30-May-1997
C;Accession: E47719
R;Wedderburn, L.R.; O'Hehir, R.E.; Hewitt, C.R.; Lamb, J.R.; Owen, M.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 8214-8218, 1993
A;Title: In vivo clonal dominance and limited T-cell receptor usage in human CD4+ T-cell
A;Reference number: A47719; MUID:93376774; PMID:8367485
A;Accession: E47719
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-29 <WED>
A;Note: sequence extracted from NCBI backbone (NCBI:137825, NCBI:137831)
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 6.2%; Score 5; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLLVL 6
|||||
Db 21 RLLVL 25

RESULT 10

S58390

T-cell receptor beta-chain Vb6-Jb2.5 - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 23-Jul-1999
C;Accession: S58390
R;Johnston, S.L.; Straubeauch, M.; Sarkar, G.; Wettstein, P.J.

Nucleic Acids Res. 23, 3074-3075, 1995

A;Title: A novel method for sequencing members of multi-gene families.
A;Reference number: S58384; MUID:95388532; PMID:7659534
A;Accession: S58390
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-29 <JOH>
A;Cross-references: EMBL:U20304; NID:G663131; PIDN:AAAG2251.1; PID:G663132
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995
A;Note: only a part of the coding sequence is given
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 6.2%; Score 5; DB 2; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLLVL 6
|||||
Db 25 RLLVL 29

RESULT 11

PH0236

T-cell receptor Vb CDR3, Ctr2 TCR Vb12 CDR 3aa.sbt - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 10-Mar-1994 #sequence_revision 28-Oct-1994 #text_change 09-May-1997
C;Accession: PH0236
R;Hara, H.; Morita, M.; Iwaki, T.; Hatae, Y.; Kitamoto, T.; Akizuki, S.; Goto, submitted to JIPID, June 1994
A;Description: Detection of HTLV-I proviral DNA and analysis of T cell receptor Vb CDR3
A;Reference number: PH0227
A;Accession: PH0236
A;Molecule type: mRNA
A;Residues: 1-31 <HAR>
A;Experimental source: spinal cord
C;Genetics:
A;Map position: 7
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: receptor

Query Match 6.2%; Score 5; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLLVL 6
|||||
Db 22 RLLVL 26

RESULT 12

I52083

major acute phase protein - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 20-Aug-1999
C;Accession: I52083
R;Anderson, K.P.; Martin, A.D.; Heath, E.C.
Arch. Biochem. Biophys. 233, 624-635, 1984
A;Title: Rat major acute-phase protein: Biosynthesis and characterization of a cDNA clone
A;Reference number: I52083; MUID:85021400; PMID:6207775
A;Accession: I52083
A;Status: preliminary; translated from GB/EMBL/DBDJB
A;Molecule type: mRNA
A;Residues: 1-33 <RES>
A;Cross-references: GB:M26758; NID:G205301; PIDN:AAAA1567.1; PID:G205302
C;Superfamily: kininogen; cystatin homology

Query Match 6.2%; Score 5; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 12 ILLLC 16
|||||

Db 2 ILLCL 6

RESULT 13

PS0128

H-2 class I histocompatibility antigen T10 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 12-Apr-1995

C:Accession: PS0128

R;Borson, K.A.; Hunt III, S.W.; Hunkapiller, T.; Sun, Y.H.; Cheroutre, H.; Nickerson, D.

J. Exp. Med. 170, 1837-1858, 1989

A>Title: Comparison of exon 5 sequences from 35 class I genes of the BALB/c mouse.

A;Reference number: PS0114; MUID:90063453; PMID:2584927

A;Accession: PS0128

A;Molecule type: DNA

A;Residues: 1-35 <BRO>

A;Experimental source: strain BALB/c

A;Note: This sequence is encoded by exon 5

C;Comment: This protein is a surface glycoprotein noncovalently associated with beta 2-m

C;Superfamily: class I histocompatibility antigen; immunoglobulin homology

C;Keywords: glycoprotein; transmembrane protein; transplantation antigen

F;1-5/Domain: connecting peptide #status predicted <COP>

F;10-30/Domain: transmembrane #status predicted <TRS>

F;31-35/Domain: intracellular #status predicted <INT>

Query Match 6.2%; Score 5; DB 2; Length 35;

Best Local Similarity 100.0%; Pred. No. 4.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 VVPGA 77

Db 16 VVPGA 20

RESULT 14

B81140

50S ribosomal protein L36 NMB0941 [imported] - Neisseria meningitidis (strain MC58 serog

C:Species: Neisseria meningitidis

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001

C:Accession: B81140; D81880

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Massignani, V.; Pizza, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve

A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: B81000; MUID:20175755; PMID:10710307

A;Accession: B81140

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-41 <TET>

A;Cross-references: GB:AE002445; GB:AE002098; NID:97226173; PIDN:AAF41347.1; PID:9722617

A;Experimental source: serogroup B, strain MC58

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:2022556; PMID:10761919

A;Accession: B81880

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-41 <PAR>

A;Cross-references: GB:AL162755; GB:AL157959; NID:97379742; PIDN:CAB84399.1; PID:9737983

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: rpmJ2; NMB0941; NMA1137

C;Superfamily: Escherichia coli ribosomal protein L36

Query Match

Best Local Similarity 6.2%; Score 5; DB 2; Length 41;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VLSSL 9

Db 3 VLSSL 7

RESULT 15

D82269

probable ribosomal protein L36 VC0879 [imported] - Vibrio cholerae (strain N16961 serog

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: D82269

R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.

chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A;Reference number: A82035; MUID:20406833; PMID:10952301

A;Accession: D82269

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-41 <HEI>

A;Cross-references: GB:AE004172; GB:AE003852; NID:99655323; PIDN:AAF94041.1; GSPDB:GN00

A;Experimental source: serogroup O1; strain N16961; biotype El Tor

C;Genetics:

A;Gene: VC0879

A;Map position: 1

C;Superfamily: Escherichia coli ribosomal protein L36

Query Match

Best Local Similarity 6.2%; Score 5; DB 2; Length 41;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VLSSL 9

Db 3 VLSSL 7

Search completed: May 16, 2003, 15:35:27

Job time : 18 secs

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=CV. BOMI RISO 1508;
 RX MEDLINE=92037994; PubMed=1936240;
 RA Hejgaard J., Jacobsen S., Svendsen I.;
 RT "Two antifungal thaumatin-like proteins from barley grain.";
 RL FEBS Lett. 231:127-131(1991).
 CC -1- FUNCTION: HAS ANTIFUNGAL ACTIVITY. INHIBITS THE GROWTH OF
 CC TRICHODERMA VIRIDAE AND CANDIDA ALBICANS.
 CC -1- SIMILARITY: BELONGS TO THE THAUMATIN FAMILY.
 CC PIR: S17573; S17573.
 DR HSP; P25871; 1AUN.
 DR InterPro: IPR001938; Thaumatin.
 DR Pfam: PF00314; thaumatin; 1.
 DR ProDom: PD001321; Thaumatin; 1.
 DR PROSITE; PS00316; THAUMATIN; PARTIAL.
 KW Plant defense; Fungicide.
 FT NON TER 44 44
 SQ SEQUENCE 44 AA; 4453 MW; 920713BCE551C1E4 CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 44;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 75 PGALP 79
 Db 15 PGALP 19
 RESULT 3
 ECNB_CITR STANDARD; PRT; 48 AA.
 AC P56550;
 DT 15-JUL-1998 (Rel. 36, Created).
 DT 15-JUL-1998 (Rel. 36, Last sequence update).
 DT 15-DEC-1998 (Rel. 37, Last annotation update).
 DE Entericidin B precursor.
 GN ECNB.
 OS Citrobacter freundii.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Citrobacter.
 OX NCBI_TaxID=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OS60;
 RX MEDLINE=98344100; PubMed=9677290;
 RA Bishop R.E., Leski B.K., Hodges R.S., Kay C.M., Weiner J.H.;
 RT "The entericidin locus of Escherichia coli and its implications for
 programmed bacterial cell death.";
 RL J. Mol. Biol. 280:593-596(1998).
 CC -1- FUNCTION: PLAYS A ROLE IN THE BACTERIOLYSIS. IS ACTIVATED UNDER
 CC CONDITIONS OF HIGH OSMOLARITY BY THE FACTOR SIGMA S. ENTERICIDIN A
 CC FUNCTIONS AS A ANTIDOTE.
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor.
 CC -1- SIMILARITY: STRONG, TO ECNA.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U21727; AAC46459.1; -.
 DR MEMBRANE; Lipoprotein; Signal.
 KW SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 48 ENTERICIDIN B.
 FT LIPID 22 22 N-ACYL DIGLYCERIDE (PROBABLE).
 SQ SEQUENCE 48 AA; 4767 MW; D951A3C8236FD18E CRC64;

Query Match 6.2%; Score 5; DB 1; Length 48;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 LVLSS 8
 Db 13 LVLSS 17
 RESULT 4
 NUSM_ARTSA STANDARD; PRT; 59 AA.
 ID NUSM_ARTSA
 AC P19047;
 DT 01-NOV-1990 (Rel. 16, Created).
 DT 01-NOV-1990 (Rel. 16, Last sequence update).
 DT 01-FEB-1994 (Rel. 28, Last annotation update).
 DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3) (Fragment).
 GN ND5.
 OS Artemia salina (Brine shrimp).
 OG Mitochondrion.
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
 OC Branchiopoda; Anostraca; Artemiidae; Artemia.
 OX NCBI_TaxID=85549;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88289417; PubMed=3135541;
 RA Batuecas B., Garesse R., Calleja M., Valverde J.R., Marco R.;
 RT "Genome organization of Artemia mitochondrial DNA.";
 RL Nucleic Acids Res. 16:6515-6529(1988).
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X07663; CAA30510.1; -.
 DR PIR; S01877; S01877.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 OC NON TER 1
 FT NON TER 1
 SQ SEQUENCE 59 AA; 6585 MW; 60C9E6E04827DC23 CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 59;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 6 LSSLL 10
 Db 46 LSSLL 50
 RESULT 5
 Y083_ARCFU STANDARD; PRT; 61 AA.
 ID Y083_ARCFU
 AC O30153;
 DT 16-OCT-2001 (Rel. 40, Created).
 DT 16-OCT-2001 (Rel. 40, Last sequence update).
 DT 16-OCT-2001 (Rel. 40, Last annotation update).
 DE Hypothetical protein AF0083.
 GN AF0083.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,

RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
 RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyridides N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,
 RA Corton M.D., Spriggs T., Artach P., Kaine B.P., Sykes S.M.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon *Archaeoglobus fulgidus*.";
 RL Nature 390:364-370(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE001100; AAB91149.1; -;
 DR TIGR; AF0083; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 61 AA; 7053 MW; 4096E1B2A28E7D36 CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 61;
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 RPAKA 31
 Db 43 RPAKA 47
 RESULT 6
 ID CXA2 CONTE STANDARD; PRT; 66 AA.
 AC Q9XZK7;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Alpha-type conotoxin Tx2 precursor.
 OS Conus textile (Cloth-of-gold cone).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Caenogastropoda;
 OC Neogastropoda; Conoidea; Conidae; Conus.
 OX NCBI_TaxID=6494;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Venom duct;
 RX MEDLINE=20037955; PubMed=10573284;
 RA Lu B.-S., Yu P., Zhao D., Huang P.-T., Huang C.-P.;
 RT "Conopeptides from *Conus striatus* and *Conus textile* by cDNA
 RT cloning.";
 RL Peptides 20:1139-1144(1999).
 CC -!- FUNCTION: Alpha-conotoxins act on postsynaptic membranes, they
 CC bind to the nicotinic acetylcholine receptors (nAChR) and thus
 CC inhibit them (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
 CC -!- SIMILARITY: BELONGS TO THE A-SUPERFAMILY OF CONOTOXINS. ALPHA-TYPE
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF146353; AAD31913.1; -;

KW Postsynaptic neurotoxin; Neurotoxin; Toxin;
 KW Acetylcholine receptor inhibitor; Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT POPEP 22 48 POTENTIAL.
 FT PRPTIDE 49 66 ALPHA-TYPE CONOTOXIN TX2.
 FT DISULFID 51 57 BY SIMILARITY.
 FT SESULEFID 52 65 BY SIMILARITY.
 SQ SEQUENCE 66 AA; 7254 MW; EDD59BBAB94F26F CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 66;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 33 SGRRT 37
 Db 24 SGRRT 28
 RESULT 7
 ID VC18 SPVKA STANDARD; PRT; 67 AA.
 AC P32217;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE Hypothetical protein C18.
 GN C18.
 OS Swinepox virus (strain Kasza) (SPV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Suipoxvirus.
 OX NCBI_TaxID=10277;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94069924; PubMed=8249275;
 RA Massung R.F., Jayarama V., Moyer R.W.;
 RT "DNA sequence analysis of conserved and unique regions of swinepox
 RT virus: identification of genetic elements supporting phenotypic
 RT observations including a novel G protein-coupled receptor
 RT homologue.";
 RL Virology 197:511-528(1993).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L22013; AAC37853.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 67 AA; 8258 MW; 4D34A27408B74975 CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 67;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 10 LCILL 14
 Db 16 LCILL 20
 RESULT 8
 ID L2MU ADECC STANDARD; PRT; 68 AA.
 AC Q65953;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Late L2 mu core protein precursor (pMu) (Protein X).
 GN PX.
 OS Canine adenovirus type 1 (strain CLL), and
 OS Canine adenovirus type 1 (strain RI261).

RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RL L72.";
RN Nature 413:852-856(2001).
[10]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jegels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
[11]
RN SEQUENCE FROM N.A.
RP SPECIES=Y.pestis; STRAIN=CO-92 / Biovar Orientalis;
RC MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RA "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
[12]
RN MASS SPECTROMETRY.
RP SPECIES=E.coli; STRAIN=K12;
RC MEDLINE=99196679; PubMed=10094780;
RA Arnold R.J., Reilly J.P.;
RT "Observation of Escherichia coli ribosomal proteins and their
posttranslational modifications by mass spectrometry.";
RL Anal. Biochem. 269:105-112(1999)
CC -1- MASS SPECTROMETRY: MW=8368.8; METHOD=MALDI.
CC -1- SIMILARITY: BELONGS TO THE SLIP FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; J01687; AAA24599.1; -;
CC EMBL; V00346; CAA23635.1; -;
CC EMBL; U28379; AAA89145.1; -;
CC EMBL; AE000388; AAC76101.1; -;
CC EMBL; AE005535; AAG58199.1; -;
CC EMBL; AP002564; BAB37371.1; -;
CC EMBL; M14427; AAA27240.1; -;
CC EMBL; AE008847; AAL22083.1; -;
CC EMBL; AL627278; CAD07734.1; -;
CC EMBL; AJ414144; CAC89499.1; -;
CC PIR; A02749; R3EC21.
CC PIR; A23985; R3EB21.
CC EcoGene; EG10920; rpsU.
CC StyGene; SG10376; rpsU.
CC InterPro; IPR001911; Ribosomal_S21.
CC Pfam; PF01165; Ribosomal_S21; 1.
CC PRINTS; PR00976; Ribosomal_S21.
CC ProDom; PD005521; Ribosomal_S21; 1.
CC TIGRfam; TIGR00030; S21p; 1.
CC DR PROSITE; PS01181; RIBOSOMAL_S21; 1.
KW Ribosomal protein; Complete proteome.
FT INIT_MET 0

SQ SEQUENCE 70 AA; 8369 MW; 2BB7C7B2FBB29013 CRC64;
Query Match 6.2%; Score 5; DB 1; Length 70;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 35 RRTRL 39
Db 65 RRTRL 69
|||||
RESULT 10
YC47_ODOSI STANDARD; PRT; 74 AA.
ID YC47_ODOSI
AC P49542;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Hypothetical 8.4 kDa protein ycf47 (ORF74).
GN YCF47.
OS Odontella sinensis (Marine centric diatom).
OC Chloroplast.
OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;
OC Biddulphiophycidae; Eupodiscaceae; Odontella.
OX NCBI_TaxID=2839;
RN [1]
RP SEQUENCE FROM N.A.
RA Kowalik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;
RT "The chloroplast genome of a chlorophyll a+c-containing alga,
RT Odontella sinensis.";
RL Plant Mol. Biol. Rep. 13:336-342(1995).
CC -1- SIMILARITY: BELONGS TO THE YCF47 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z67753; CAA91678.1; -;
CC KX Chloroplast; Hypothetical protein.
SQ SEQUENCE 74 AA; 8436 MW; A80A858F7A939C72 CRC64;
Query Match 6.2%; Score 5; DB 1; Length 74;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 46 SPNST 50
Db 41 SPNST 45
|||||
RESULT 11
YC55_HAEIN STANDARD; PRT; 74 AA.
ID YC55_HAEIN
AC P44121;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein H11255.
GN H11255.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
 Rd";
 RL Science 269:496-512(1995).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U32805; AAC22910.1; --
 DR TIGR; H11255; --
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 74 AA; 8446 MW; CC22E0A66C0C98D3 CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 74;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 24 GKRRP 28
 DB 13 GKRRP 17
 |||||
 RESULT 12
 GRFA SFVKA
 ID GRFA SFVKA STANDARD; PRT; 80 AA.
 AC P08411;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE Growth factor.
 OS Shope fibroma virus (strain Kasza) (SFV).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Leporipoxvirus.
 OX NCBI_TaxID=10272;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87172751; PubMed=3031480;
 RA Chang W., Upton C., Hu S.-L., Purchio A.F., McFadden G.;
 RT "The genome of Shope fibroma virus, a tumorigenic poxvirus, contains
 RT a growth factor gene with sequence similarity to those encoding
 RT epidermal growth factor and transforming growth factor alpha.";
 RL Mol. Cell. Biol. 7:535-540(1987).
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M15921; AAA66873.1; --
 DR PIR; A26723; EGVZSF.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001336; EGF_1.
 DR Pfam; PF00008; EGF_1.
 DR PRINTS; PR00009; EGFTGF.
 DR SMART; SM00181; EGF_1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; FALSE_NEG.
 KW EGF-like domain; Growth factor; Glycoprotein.

FT DOMAIN 29 73 EGF-LIKE.
 FT DISULFID 33 47 BY SIMILARITY.
 FT DISULFID 41 61 BY SIMILARITY.
 FT DISULFID 63 72 BY SIMILARITY.
 FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 80 AA; 9210 MW; C48D30E878D2ED58 CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 8 SLLCI 12
 DB 9 SLLCI 13
 |||||
 RESULT 13
 ICTA SYN3
 ID ICTA SYN3 STANDARD; PRT; 80 AA.
 AC P27372;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Inorganic carbon transport protein.
 GN ICTA OR SSRI386.
 OS Synechocystis sp. (strain PCC 6803).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ogawa T.;
 RT "Cloning and inactivation of a gene essential to inorganic carbon
 RT transport of Synechocystis PCC6803.";
 RL Plant Physiol. 96:280-284(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M73833; AAA27290.1; --
 DR EMBL; D90911; BAA18123.1; --
 DR PIR; JQ1959; JQ1959.
 KW Transport; Complete proteome.
 SQ SEQUENCE 80 AA; 9252 MW; 311AD10FA3B6136D CRC64;
 Query Match 6.2%; Score 5; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 LLVLS 7
 DB 64 LLVLS 68
 |||||
 RESULT 14
 ACHG_ELEEL

ID ACHG ELSEEL STANDARD; PRT; 23 AA.
 AC P09692;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acetylcholine receptor protein, gamma chain (Fragment).
 OS Electrophorus electricus (Electric eel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Gymnotiformes;
 OC Electrophoridae; Electrophorus.
 OX NCBI_TaxID=8005;
 RN [1]
 RP SEQUENCE
 RA MEDLINE=83065199; PubMed=6959131;
 RA Conti-Troconci B.M., Hunkapillier M.W., Lindstrom J.M., Raftery M.A.;
 RT "Subunit structure of the acetylcholine receptor from Electrophorus
 RT electricus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 79:6489-6493 (1982).
 CC -!- FUNCTION: AFTER BINDING ACETYLCHOLINE, THE ACHR RESPONDS BY AN
 CC EXTENSIVE CHANGE IN CONFORMATION THAT AFFECTS ALL SUBUNITS AND
 CC LEADS TO OPENING OF AN ION-CONDUCTING CHANNEL ACROSS THE PLASMA
 CC MEMBRANE.
 CC -!- SUBUNIT: PENTAMER OF TWO ALPHA CHAINS, AND ONE EACH OF THE BETA,
 CC DELTA, AND GAMMA CHAINS.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
 DR PIR; C27262; C27262.
 DR InterPro; IPR001175; Neur channel.
 DR Pfam; PF02931; Neur chan LBD; 1.
 DR PROSITE; PS00236; NEUROTR ION_CHANNEL; PARTIAL.
 KW Receptor; Postsynaptic membrane; Ionic channel; Transmembrane.
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2717 MW; DA88F7E990C024B CRC64;

 Query Match 4.9%; Score 4; DB 1; Length 23;
 Best Local Similarity 100.0%; Pred.No.1.1e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 27 RPAK 30
 Db 19 RPAK 22

 RESULT 15
 ID UHAA4 HUMAN STANDARD; PRT; 23 AA.
 AC P49289;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of heart (Fragments).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Heart;
 RX MEDLINE=96007936; PubMed=7498159;
 RA Kovalyov L.I., Shishkin S.S., Efimochkin A.S., Kovalyova M.A.,
 RA Ershova E.S., Egorov T.A., Musalyamov A.K.;
 RT "The major protein expression profile and two-dimensional protein
 RT database of human heart.";
 RL Electrophoresis 16:1160-1169 (1995).
 CC -!- CAUTION: WAS ORIGINALLY (REF.1) THOUGHT TO BE MITOCHONDRIAL
 CC ACONITASE.
 FT NON_TER 1 1
 FT NON_CONS 16 17
 FT NON_TER 23 23
 SQ SEQUENCE 23 AA; 2388 MW; 90EF513CEEB8B451 CRC64;

 Query Match 4.9%; Score 4; DB 1; Length 23;
 Best Local Similarity 100.0%; Pred.No.1.1e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 50 TNLK 53
 Db 20 TNLK 23
 Search completed: May 16, 2003, 15:34:26
 Job time : 14 secs

THIS PAGE BLANK (UBPTD)

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:31:20 ; Search time 30 Seconds
(without alignments)
556.327 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLVLSLLCILLCSIF.....PCKLEPRLWVVGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 70278

Minimum DB seq length: 23

Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database :

SPTREMBL 21:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mic:.*
8: sp_organelle:.*
9: sp_phase:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_rvirus:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7	8.6	51	5	O61655 ceratitis c
2	7	8.6	73	16	O8XRN8 ralatonia s
3	6	7.4	30	4	Q86D69 homo sapien
4	6	7.4	52	10	Q9SK17 arabidopsis
5	6	7.4	73	4	Q8BT96 homo sapien
6	6	7.4	76	10	Q9XER1 gossypium h
7	6	7.4	77	3	Q9P5V0 neurospora
8	6	7.4	78	16	Q9ALT5 streptococ
9	6	7.4	79	12	Q9ILC1 white spot
10	6	7.4	79	16	Q9PD22 xylella fas
11	5	6.2	23	10	Q9S906 triticum ae
12	5	6.2	24	8	Q94XE2 tectocoris
13	5	6.2	25	11	Q9QUX1 rattus sp.
14	5	6.2	25	13	Q9PWR9 xiphophorus
15	5	6.2	26	10	Q9S907 avena sativ
16	5	6.2	27	12	Q9Q9F2 hepatitis c

17	5	6.2	27	12	Q9Q9E6	Q9q9e6 hepatitis c
18	5	6.2	27	12	Q9Q9E2	Q9q9e2 hepatitis c
19	5	6.2	27	12	Q9Q9D6	Q9q9d6 hepatitis c
20	5	6.2	28	4	O43804	O43804 homo sapien
21	5	6.2	28	8	O98676	O98676 ondinea pur
22	5	6.2	29	10	Q9S8C7	Q9s8c7 secale cere
23	5	6.2	30	7	Q31234	Q31234 mus musculu
24	5	6.2	30	11	Q9WVD6	Q9wvd6 mus musculu
25	5	6.2	31	2	O47323	O47323 escherichia
26	5	6.2	32	7	Q9BCT8	Q9bct8 homo sapien
27	5	6.2	32	7	Q95HL3	Q95hl3 homo sapien
28	5	6.2	32	7	Q8SNB9	Q8snb9 homo sapien
29	5	6.2	32	11	Q9RIC2	Q9ric2 mus musculu
30	5	6.2	33	11	O63329	O63329 rattus norv
31	5	6.2	34	4	Q96J93	Q96j93 homo sapien
32	5	6.2	35	2	Q9R4S7	Q9r4s7 rhodobacter
33	5	6.2	35	4	Q9NOY7	Q9nvy7 homo sapien
34	5	6.2	35	4	Q9UGE7	Q9uge7 homo sapien
35	5	6.2	37	2	Q9EY28	Q9ey28 photobacter
36	5	6.2	37	4	Q9NQN5	Q9eqn5 homo sapien
37	5	6.2	37	5	Q9BM33	Q9bm33 themiste al
38	5	6.2	37	5	Q9BM32	Q9bm32 themiste al
39	5	6.2	39	3	O8TGR6	O8tgr6 saccharomyc
40	5	6.2	39	12	O88527	O88527 turkey herp
41	5	6.2	39	16	O8XTC6	O8xtc6 ralatonia s
42	5	6.2	40	12	O91X09	O91x09 hepatitis c
43	5	6.2	41	15	Q9PXE3	Q9pxe3 chimpanzee
44	5	6.2	41	16	Q9PAQ5	Q9paq5 xylella fas
45	5	6.2	41	16	Q9KTM3	Q9krm3 vibrio chol

ALIGNMENTS

RESULT 1

O61655 PRELIMINARY; PRT; 51 AA.
AC O61655;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome P450 monooxygenase (Fragment).
OS Ceratitis capitata (Mediterranean fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Tephritidae; Tephritidae; Ceratitis.
OX NCBI_TaxID=7213;
RN [1]
RP SEQUENCE FROM N.A.
RA Danielson P.B., Foster J.L.M., Cooper S.K., Fogleman J.C.;
RT "Diversity of Expressed cytochrome P450 Genes in the Adult
RT Mediterranean Fruit Fly, Ceratitis capitata."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF056474; AAC13307.1; -
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; p450; 1.
KW Monooxygenase.
FT NON TER 1 1
SQ SEQUENCE 51 AA; 5854 MW; 0CAlF3B75BDECE44 CRC64;

Query Match 8.6%; Score 7; DB 5; Length 51;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 4 LVLSSLL 10
Db 10 LVLSSLL 16

RESULT 2

O8XRN8 PRELIMINARY; PRT; 73 AA.
ID O8XRN8
AC O8XRN8;

```

DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein RSP0793.
GN RSP0793 OR RS01914.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GMI1000;
RX MEDLINE=21681879; PubMed=11823852;
RA Salancoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:457-502(2002).
DR EMBL; AL646080; CAD17944.1; -.
KW Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 73 AA; 8333 MW; 8BEFAAC58DAE173C CRC64;

Query Match 8.6%; Score 7; DB 16; Length 73;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 EGKRRPA 29
Db 29 EGKRRPA 35
|||||

RESULT 3
Q96D69 PRELIMINARY; PRT; 30 AA.
AC Q96D69;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical 3.1 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC012777; AAH12777.1; -.
KW Hypothetical protein.
SQ SEQUENCE 30 AA; 3108 MW; C5EB62C1A5F8405C CRC64;

Query Match 7.4%; Score 6; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 73 VVPGAL 78
Db 15 VVPGAL 20
|||||

RESULT 4
Q9SK17 PRELIMINARY; PRT; 52 AA.
AC Q9SK17;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AT2g06480 protein.

```

```

GN AT2G06480.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsiis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carreia A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salzberg S.B., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006420; AAD25147.1; -.
SQ SEQUENCE 52 AA; 5985 MW; 82F906B7FB4DCEE1 CRC64;

Query Match 7.4%; Score 6; DB 10; Length 52;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 VPGALP 79
Db 2 VPGALP 7
|||||

RESULT 5
Q9BT96 PRELIMINARY; PRT; 73 AA.
AC Q9BT96;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein FLJ10719.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SKIN;
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004277; AAH04277.1; -.
KW Hypothetical protein.
SQ SEQUENCE 73 AA; 8217 MW; 64A714046E960DD CRC64;

Query Match 7.4%; Score 6; DB 4; Length 73;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LVLVSS 8
Db 38 LVLVSS 43
|||||

RESULT 6
Q9XER1 PRELIMINARY; PRT; 76 AA.
AC Q9XER1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)

```



```

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE FS18A precursor.
GN Gossypium hirsutum (Upland cotton).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Malvales; Malvaceae; Gossypium.
OX NCBI_TaxID=3635;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. SIOKA 1-4;
RA Orford S.J., Carney T.J., Olesnick N.S., Timmis J.N.;
RT "Characterization of a cotton gene expressed late in fiber cell
RT elongation.";
RL Theor. Appl. Genet. 0:0-0(1999).
DR EMBL; AF114254; AAD31394.1; -.
KW SIGNAL.
FT CHAIN 1 26 POTENTIAL.
FT SIGNAL 27 76 PROCESSED FS18A.
SQ SEQUENCE 76 AA; 7556 MW; 0C3FB299AE7DB57 CRC64;

Query Match 7.4%; Score 6; DB 10; Length 76;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLLC 11
Db 68 LSSLLC 73

RESULT 7
Q9PSV0 PRELIMINARY; PRT; 77 AA.
AC Q9PSV0;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Conserved hypothetical protein.
GN B5022.20.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL355932; CAB91421.1; -.
KW Hypothetical protein.
SQ SEQUENCE 77 AA; 8846 MW; D6FD4B9CB299A7E3 CRC64;

Query Match 7.4%; Score 6; DB 3; Length 77;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LCILLL 15
Db 62 LCILLL 67

RESULT 8
Q9A1T5 PRELIMINARY; PRT; 78 AA.
AC Q9A1T5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Hypothetical protein SPY0105.

```

```

GN SPY0105.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvoarov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006481; AAK33223.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 78 AA; 8898 MW; 6EC41E2909801EE1 CRC64;

Query Match 7.4%; Score 6; DB 16; Length 78;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSSL 9
Db 13 LVLSSL 18

RESULT 9
Q91LC1 PRELIMINARY; PRT; 79 AA.
AC Q91LC1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE ORF124 (Wsv247) (WSSV302).
OS White spot syndrome virus (WSSV).
OC Viruses; unclassified viruses.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21342572; PubMed=11448154;
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlask J.M.;
RT "The white spot syndrome virus DNA genome sequence.";
RL Virology 286:7-22(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA van Hulten M.C.W., Witteveldt J., Peters S., Kloosterboer N.,
RA Tarchini R., Fiers M., Sandbrink H., Lankhorst R.K., Vlask J.M.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21548311; PubMed=11689662;
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RT "Complete genome sequence of the shrimp white spot bacilliform
RT virus.";
RL J. Virol. 75:11811-11820(2001).
RN [4]
RP SEQUENCE FROM N.A.
RA Yang F., He J., Lin X., Li Q., Pan D., Zhang X., Xu X.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=TAIWAN;
RX MEDLINE=20517548; PubMed=11062040;
RA Tsai M.F., Yu H.T., Tzeng H.F., Leu J.H., Chou C.M., Huang C.J.,
RA Wang C.H., Lin J.Y., Kou G.H., Lo C.F.;
RT "Identification and characterization of a shrimp white spot syndrome
RT virus (WSSV) gene that encodes a novel chimeric polypeptide of
RT cellular-type thymidine kinase and thymidylate kinase.";
RL Virology 277:100-110(2000).
RN [6]

```

RP SEQUENCE FROM N.A.
RC STRAIN=TAIWAN;
RX MEDLINE=21844071; PubMed=11853338;
RA Chen L.L., Liu J.H., Huang C.J., Chou C.M., Chen S.M., Wang C.H.,
Lo C.F., Kou G.H.;
RT "Identification of a nucleocapsid protein (VP35) gene of shrimp white
RT spot syndrome virus and characterization of the motif important for
RT targeting VP35 to the nuclei of transfected insect cells";
RL Virology 293:44-53 (2002).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=TAIWAN;
RA Lo C.-F., Kou G.-H.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF369029; AAK7793.1; -
DR EMBL; AF332093; AAL33250.1; -
DR EMBL; AF440570; AAL89170.1; -
SQ SEQUENCE 79 AA; 9285 MW; 734EE8DBB60FDA86 CRC64;

Query Match 7.4%; Score 6; DB 12; Length 79;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLSSLL 10
DB 58 VLSSLL 63

RESULT 10
Q9PD22 ID Q9PD22 PRELIMINARY; PRT; 79 AA.
AC Q9PD22, 2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Single-stranded DNA binding protein.
GN xfl558
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
Faciniani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
Machado M.A., Madeira A.M.B.L., Madeira H.M.F., Marino C.L.,
Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
Meck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa";
RL Nature 406:151-159 (2000).

DR EMBL; AE003985; AAF84367.1; -
DR HSP; P02339; 1EYC.
DR InterPro; IPR000424; SSB_protein.
DR Pfam; PF00436; SSB; 1.
KW Complete proteome.
SQ SEQUENCE 79 AA; 8935 MW; 2DEEA09C35D4EADD CRC64;

Query Match 7.4%; Score 6; DB 16; Length 79;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 RRPAAK 31
DB 53 RRPAAK 58

RESULT 11
Q9S906 ID Q9S906 PRELIMINARY; PRT; 23 AA.
AC Q9S906, 2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Antifungal protein (Fragment).
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE.
RX MEDLINE=92190627; PubMed=1799695;
RA Vigers A.J., Roberts W.K., Selitrennikoff C.P.;
RT "A new family of plant antifungal proteins";
RL Mol. Plant Microbe Interact. 4:315-323 (1991).
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; Thaumatin; 1.
FT NON_TER 1 1
FT NON_TER 23 23
SQ SEQUENCE 23 AA; 2291 MW; C3D88BA722C21D7 CRC64;

Query Match 6.2%; Score 5; DB 10; Length 23;
Best Local Similarity 100.0%; Pred. No. 4.1e-02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 75 PGALP 79
DB 15 PGALP 19

RESULT 12
Q94XE2 ID Q94XE2 PRELIMINARY; PRT; 24 AA.
AC Q94XE2, 2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NADH dehydrogenase subunit 4 (Fragment).
GN NAD4.
OS Tectocoris diophthalmus (cotton harlequin bug).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hemiptera; Euhemiptera;
OC Psyllota; Neoptera; Paraneoptera; Hemiptera; Pentatomidae;
OC Heteroptera; Pantheroptera; Pentatomomorpha; Pentatomidae;
OC Pentatomidae; Tectocoris.
OX NCBI_TaxID=159956;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21396409; PubMed=11504862;
RA Shao R., Campbell N.J.H., Schmidt E.R., Barker S.C.;
RT "Increased rate of gene rearrangement in the mitochondrial genomes of
RT three orders of hemipteroid insects";
RL Mol. Biol. Evol. 18:1828-1832 (2001).

DR EMBL; AF335991; AAK55287.1; --

KW Mitochondrion. 1

FT NON_TER 1

SQ SEQUENCE 24 AA; 2871 MW; 972B55027DC29675 CRC64;

Query Match

6.2%; Score 5; DB 8; Length 24;

Best Local Similarity 100.0%; Pred. No. 4.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 17 FSIFS 21

|||||

Db 2 FSIFS 6

RESULT 13

O9QUX1

ID Q9QUX1 PRELIMINARY; PRT; 25 AA.

AC Q9QUX1;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

DE N-acetyl-beta-D-hexosaminidase (EC 3.2.1.52) (Fragment).

OS Rattus sp.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10118;

RN [1]

RP SEQUENCE.

RX MEDLINE=96249836; PubMed=8924513;

RA Hall J.C., Perez F.M., Kochins J.G., Pettersen C.A., Li Y.,

RA Tubbs C.E., LaMarche M.D.;

RT "Quantification and localization of N-acetyl-beta-D-hexosaminidase in

the adult rat testis and epididymis."

RL Biol. Reprod. 54:914-929(1996).

SQ SEQUENCE 25 AA; 2722 MW; CBC97D5B4ED97CA0 CRC64;

Query Match

6.2%; Score 5; DB 11; Length 25;

Best Local Similarity 100.0%; Pred. No. 4.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 46 SPNST 50

|||||

Db 17 SPNST 21

RESULT 14

O9PWR9

ID O9PWR9 PRELIMINARY; PRT; 25 AA.

AC O9PWR9;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Melanoma receptor tyrosine kinase (Fragment).

OS Xiphophorus maculatus (Southern platyfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;

OC Cyprinodontiformes; Poeciliidae; Xiphophorus.

OX NCBI_TaxID=8083;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ONC SR-STRAIN;

RX MEDLINE=99132631; PubMed=9931413;

RA Scharlt M., Wilde B., Hornung U.;

RT "Triplet repeat variability in the signal peptide sequence of the Xmrk

receptor tyrosine kinase gene in Xiphophorus fish."

RL Gene 224:17-21(1998).

DR EMBL; U82804; AAD10123.1; --

FT NON_TER 25

KW Kinase; Receptor.

SQ SEQUENCE 25 AA; 2695 MW; 967AF74362DF4350 CRC64;

Query Match

6.2%; Score 5; DB 13; Length 25;

Best Local Similarity

100.0%; Pred. No. 4.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 LVLVS 7

|||||

Db 14 LVLVS 18

RESULT 15

O9S907

ID O9S907 PRELIMINARY; PRT; 26 AA.

AC O9S907;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Antifungal protein (Fragment).

OS Avena sativa (Oat).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Aveneae; Avena.

OX NCBI_TaxID=4498;

RN [1]

RP SEQUENCE.

RX MEDLINE=92190627; PubMed=1799695;

RA Vigers A.J., Roberts W.K., Seiltrennikoff C.P.;

RT "A new family of plant antifungal proteins."

RL Mol. Plant Microbe Interact. 4:315-323(1991).

DR InterPro; IPR001938; Thaumatin.

DR Pfam; PF00314; Thaumatin.1.

FT NON_TER 1

FT NON_TER 26

SQ SEQUENCE 26 AA; 2648 MW; 57C46A9E388E4A72 CRC64;

Query Match

6.2%; Score 5; DB 10; Length 26;

Best Local Similarity 100.0%; Pred. No. 4.6e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 75 PGALP 79

|||||

Db 15 PGALP 19

Search completed: May 16, 2003, 15:35:05

Job time : 33 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:34:30 ; Search time 29 Seconds
(without alignments)
82.181 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 81
Sequence: 1 MRLVLVLSLLCILLCSIF.....PCKLEPEPRLWVVGALPQV 81

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 51475

Minimum DB seq length: 23

Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5	6.2	23	1	US-08-086-428B-159
2	5	6.2	23	2	US-08-468-570-159
3	5	6.2	23	2	US-08-290-665A-263
4	5	6.2	23	4	US-08-469-260A-400
5	5	6.2	23	5	PCT-US95-10398-263
6	5	6.2	24	1	US-07-928-930A-7
7	5	6.2	24	1	US-08-288-568-7
8	5	6.2	24	1	US-08-487-461-7
9	5	6.2	24	1	US-08-432-691-7
10	5	6.2	24	1	US-08-487-459-7
11	5	6.2	24	4	US-09-288-143-127
12	5	6.2	27	1	US-08-066-325-89
13	5	6.2	27	4	US-08-066-325-89
14	5	6.2	27	4	US-08-965-762-8
15	5	6.2	27	4	US-08-911-927-8
16	5	6.2	28	4	US-09-507-819-30
17	5	6.2	29	1	US-08-066-325-61
18	5	6.2	29	1	US-08-066-325-74
19	5	6.2	29	1	US-08-066-325-75
20	5	6.2	29	2	US-08-419-061A-5
21	5	6.2	29	2	US-08-485-647A-5
22	5	6.2	29	4	US-09-507-819-46
23	5	6.2	29	4	US-09-028-274A-5
24	5	6.2	30	1	US-08-066-325-60
25	5	6.2	30	4	US-08-043-293-1
26	5	6.2	30	4	US-09-511-023-5
27	5	6.2	31	1	US-08-066-325-76

28	5	6.2	31	1	US-08-066-325-77	Sequence 77, Appl
29	5	6.2	31	1	US-08-066-325-78	Sequence 78, Appl
30	5	6.2	31	4	US-09-507-819-48	Sequence 48, Appl
31	5	6.2	32	4	US-09-461-697-157	Sequence 157, App
32	5	6.2	32	4	US-09-227-357-207	Sequence 207, App
33	5	6.2	34	1	US-07-915-247A-21	Sequence 21, Appl
34	5	6.2	34	1	US-08-443-863-21	Sequence 21, Appl
35	5	6.2	34	1	US-08-448-070-21	Sequence 21, Appl
36	5	6.2	34	1	US-08-449-500-21	Sequence 21, Appl
37	5	6.2	34	1	US-08-449-317A-21	Sequence 21, Appl
38	5	6.2	34	2	US-08-477-022-21	Sequence 21, Appl
39	5	6.2	34	2	US-08-449-447-21	Sequence 21, Appl
40	5	6.2	34	2	US-08-184-328-21	Sequence 21, Appl
41	5	6.2	34	2	US-08-521-097-21	Sequence 21, Appl
42	5	6.2	35	1	US-08-449-500-84	Sequence 84, Appl
43	5	6.2	35	1	US-08-449-317A-84	Sequence 84, Appl
44	5	6.2	35	2	US-08-477-022-84	Sequence 84, Appl
45	5	6.2	35	2	US-08-449-447-84	Sequence 84, Appl

ALIGNMENTS

RESULT 1
US-08-086-428B-159
; Sequence 159, Application US/08086428B
; Patent No. 5514539
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R. H. AND
; APPLICANT: PURCELL, R. H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,428B
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 159:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; US-08-086-428B-159

Query Match 6.2%; Score 5; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSS 8

Db 4 LVLSS 8
|||||

RESULT 2

US-08-468-570-159
; Sequence 159, Application US/08468570
; Patent No. 5871962
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 GENE
; TITLE OF INVENTION: OF 51 ISOLATES OF HEPATITIS C AND THE USE
; TITLE OF INVENTION: OF REAGENTS DERIVED FROM THESE SEQUENCES IN
; TITLE OF INVENTION: DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 159
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,570
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29-JUN-1993
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4070US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 159:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
US-08-468-570-159

Query Match 6.2%; Score 5; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSS 8
|||||
Db 4 LVLSS 8

RESULT 3

US-08-230-665A-263
; Sequence 263, Application US/08290665A
; Patent No. 5882852
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE

; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/290,665A
; FILING DATE: 15-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 263:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
US-08-290-665A-263

Query Match 6.2%; Score 5; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSS 8
|||||
Db 4 LVLSS 8

RESULT 4

US-08-469-260A-400
; Sequence 400, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUERHOFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BULJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,260A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 400:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-260A-400

Query Match 6.2%; Score 5; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 24 GKRRP 28
Db 8 GKRRP 12

RESULT 5
PCT-US95-10398-263
; Sequence 263, Application PC/TUS9510398
; GENERAL INFORMATION:
; APPLICANT: BUKH, J., MILLER, R.H. AND
; APPLICANT: PURCELL, R.H.
; TITLE OF INVENTION: NUCLEOTIDE AND DEDUCED
; TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
; TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
; TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
; TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
; NUMBER OF SEQUENCES: 263
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10398
; FILING DATE: 15-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/086,428
; FILING DATE: 29 JUNE 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/290/665
; FILING DATE: 15 AUGUST 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: RICHARD W. BORK
; REGISTRATION NUMBER: 36,459
; REFERENCE/DOCKET NUMBER: 2026-4116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 263:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; PCT-US95-10398-263

Query Match 6.2%; Score 5; DB 5; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSS 8
Db 4 LVLSS 8

RESULT 6
US-07-928-930A-7
; Sequence 7, Application US/07928930A
; Patent No. 5344822
; GENERAL INFORMATION:
; APPLICANT: Levine, Daniel M.; Parker, Thomas S.;
; APPLICANT: Rubin, Albert L.
; TITLE OF INVENTION: Methods Useful in Endotoxin
; TITLE OF INVENTION: Prophylaxis and Therapy
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,930A
; FILING DATE: 19920812
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5344822man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; US-07-928-930A-7

Query Match 6.2%; Score 5; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLL 10
Db 5 LSSLL 9

RESULT 7
US-08-288-568-7
; Sequence 7, Application US/08288568
; Patent No. 5508218
; GENERAL INFORMATION:
; APPLICANT: Parker, Thomas S.; Levine, Daniel M.;
; APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.

```

; TITLE OF INVENTION: Methods Useful in Endotoxin Based
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,568
; FILING DATE: 10-AUGUST-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07453
; FILING DATE: 9-AUGUST-1993
; NAME: Hanson, No. 5506218man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-288-568-7

Query Match 6.2%; Score 5; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLL 10
Db 5 LSSLL 9

RESULT 8
US-08-487-461-7
; Sequence 7, Application US/08487461
; Patent No. 5587366
; GENERAL INFORMATION:
; APPLICANT: Parker, Thomas S.; Levine, Daniel M.;
; APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.
; TITLE OF INVENTION: Methods Useful in Endotoxin Based
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/288,568
; FILING DATE: 10-AUGUST-1994
; NAME: Hanson, No. 5614507man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200

```

```

; APPLICATION NUMBER: US/08/487,461
; FILING DATE: June 7, 1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07453
; FILING DATE: 9-AUGUST-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,930
; FILING DATE: 12-AUGUST-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5587366man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 211.2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 24 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-487-461-7

Query Match 6.2%; Score 5; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLL 10
Db 5 LSSLL 9

RESULT 9
US-08-432-691-7
; Sequence 7, Application US/08432691
; Patent No. 5614507
; GENERAL INFORMATION:
; APPLICANT: Parker, Thomas S.; Levine, Daniel M.;
; APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.
; TITLE OF INVENTION: Methods Useful in Endotoxin Based
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb
; MEDIUM TYPE: storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/432,691
; FILING DATE: 02-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/288,568
; FILING DATE: 10-AUGUST-1994
; APPLICATION NUMBER: PCT/US93/07453
; FILING DATE: 9-AUGUST-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,930
; FILING DATE: 12-AUGUST-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5614507man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: ROGO 211
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200

```


TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-432-691-7

Query Match 6.2%; Score 5; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLL 10
Db 5 LSSLL 9

RESULT 10

US-08-487-459-7
Sequence 7, Application US/08487459
Patent No. 5674855

GENERAL INFORMATION:

APPLICANT: Levine, Daniel M.; Parker, Thomas S.;
APPLICANT: Rubin, Albert L.; Gordon, Bruce R.; Saal, Stuart D.
TITLE OF INVENTION: Methods Useful in Endotoxin Based

TITLE OF INVENTION: Prophylaxis and Therapy

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb

MEDIUM TYPE: storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,459

FILING DATE: 7-JUNE-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/07453

FILING DATE: 9-AUGUST-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/928,930

FILING DATE: 12-AUGUST-1992

ATTORNEY/AGENT INFORMATION:

NAME: Handon, No. 5674855man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: ROGO 211.3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 24 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-487-459-7

Query Match 6.2%; Score 5; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLL 10
Db 5 LSSLL 9

RESULT 11

US-09-288-143-127
Sequence 127, Application US/09288143
Patent No. 6433139
GENERAL INFORMATION:

APPLICANT: Brewer et al.

TITLE OF INVENTION: 53 Human Secreted Proteins

FILE REFERENCE: E2018P1

CURRENT APPLICATION NUMBER: US/09/288,143

CURRENT FILING DATE: 1999-04-08

EARLIER APPLICATION NUMBER: PCT/US98/21142

EARLIER FILING DATE: 1998-10-08

EARLIER APPLICATION NUMBER: 60/061,463

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,529

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/071,498

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,527

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,536

EARLIER FILING DATE: 1997-10-09

EARLIER APPLICATION NUMBER: 60/061,532

EARLIER FILING DATE: 1997-10-09

NUMBER OF SEQ ID NOS: 219

SOFTWARE: Patentin Ver. 2.0

SEQ ID NO 127

LENGTH: 24

TYPE: PRT

ORGANISM: Homo sapiens

US-09-288-143-127

Query Match 6.2%; Score 5; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LVLSS 8
Db 9 LVLSS 13

RESULT 12

US-08-066-325-89

Sequence 89, Application US/08066325

Patent No. 5667967

GENERAL INFORMATION:

APPLICANT: Steinman, Lawrence

APPLICANT: Oksenberg, Jorge

APPLICANT: Bernard, Claude

TITLE OF INVENTION: T-CELL RECEPTOR VARIABLE TRANSCRIPTS AS DISEASE RELATED MARKER

NUMBER OF SEQUENCES: 157

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/066,325

FILING DATE: 21-MAY-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5667967tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 690068.408C1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 89:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 27 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-066-325-89

Query Match 6.2%; Score 5; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RLLVL 6
|
|
|
|
Db 18 RLLVL 22

RESULT 13
US-08-965-762-8
; Sequence 8, Application US/08965762
; Patent No. 6280963
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrias, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 07334/062001
; CURRENT APPLICATION NUMBER: US/08/965,762
; CURRENT FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-08-965-762-8

Query Match 6.2%; Score 5; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 CILLL 15
|
|
|
|
Db 4 CILLL 8

RESULT 14
US-09-911-927-8
; Sequence 8, Application US/09911927
; Patent No. 6461826
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrias, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 06286-062003
; CURRENT APPLICATION NUMBER: US/09/911,927
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/965,762
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-09-911-927-8

Query Match 6.2%; Score 5; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 CILLL 15

Db 4 CILLL 8
|
|
|
|

RESULT 15
US-09-911-882-8
; Sequence 8, Application US/09911882
; Patent No. 6465198
; GENERAL INFORMATION:
; APPLICANT: Koltin, Yigal
; APPLICANT: Gavrias, Victoria
; TITLE OF INVENTION: ESSENTIAL FUNGAL GENES AND THEIR USE
; FILE REFERENCE: 06286-062004
; CURRENT APPLICATION NUMBER: US/09/911,882
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 08/965,762
; PRIOR FILING DATE: 1997-11-07
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
US-09-911-882-8

Query Match 6.2%; Score 5; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 CILLL 15
|
|
|
|
Db 4 CILLL 8

Search completed: May 16, 2003, 15:36:11
Job time : 38 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 17, 2003, 03:26:33 ; Search time 2182 Seconds
(without alignments)
1080.352 Million cell updates/sec

Title: US-09-724-000A-5
Perfect score: 442
Sequence: 1 MRLVLSSLICILLCSIF.....PCKLEPRLWVWPALPQV 81

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09724000/runat_16052003_145459_17144/app.query.fasta_1.263
-DB=GenEmbl -QPM=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09724000.cgn 1.1 2496 @runat_16052003_145459_17144 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

RESULT 1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	442	100.0	382	6	AX340424 Sequence
C 2	442	100.0	485	6	AX351259 Sequence
C 3	442	100.0	804	6	AX464016 Sequence
C 4	442	100.0	806	6	AX342218 Sequence
C 5	442	100.0	2063	9	AK025416 Homo sapi
C 6	400	90.5	797	6	AX027773 Sequence
C 7	400	90.5	801	6	AX027767 Sequence
C 8	228.5	51.7	744	6	AX342215 Sequence
C 9	183	41.4	744	10	AF152002 Rattus no
C 10	178	40.3	4159	6	AX342222 Sequence
C 11	178	40.3	198481	9	AC023389 Homo sapi
C 12	160.5	36.3	742	10	S74257 S74257 2c9 gene (c
C 13	111.5	25.2	137164	2	AC128725 Rattus no
C 14	106.5	24.1	225929	2	AC109260 Mus muscu
C 15	92	20.8	207374	9	CNS07EFU AL583722 Human chr
C 16	88.5	20.0	173235	2	AC106276 Rattus no
C 17	87.5	19.8	118429	2	AL158017 Homo sapi
C 18	87.5	19.8	198906	9	AL356275 Human DNA
C 19	86.5	19.6	2167	6	A37236 Sequence 5
C 20	86.5	19.6	2168	4	BOVPRG MI3976 Bovine gamm
C 21	86.5	19.6	108071	2	AC123218 Rattus no
C 22	86.5	19.6	161044	2	AC119102 Rattus no
C 23	86.5	19.6	176092	2	AC073252 Homo sapi
C 24	86	19.5	170000	2	AC004394 Homo sapi
C 25	86	19.5	170000	2	AC004579 Homo sapi
C 26	86	19.5	190000	2	AC004580 Homo sapi
C 27	86	19.5	210515	9	AC097382 Homo sapi
C 28	85.5	19.3	10647	1	AE002031 Deinococc
C 29	85.5	19.3	110000	2	LMFLCHR36_10 Continuation (11 o
C 30	85.5	19.3	161112	2	AC115876 Mus muscu
C 31	85	19.2	853	9	BC002624 Homo sapi
C 32	85	19.2	876	9	BC001181 Homo sapi
C 33	85	19.2	1688	6	AC035537 Sequence
C 34	85	19.2	1688	6	E11457 E11457 cDNA encodi
C 35	85	19.2	1934	9	BC007790 Homo sapi
C 36	85	19.2	1979	6	AR035536 Sequence
C 37	85	19.2	1979	6	E11456 E11456 cDNA encodi
C 38	85	19.2	2010	9	HSU07132 Human stero
C 39	85	19.2	2030	6	I36667 Sequence 1
C 40	85	19.2	2030	6	I70211 Sequence 1
C 41	85	19.2	2384	9	AK091535 Homo sapi
C 42	85	19.2	157375	2	AC122095 Rattus no
C 43	85	19.2	195983	2	AC099699 Mus muscu
C 44	84.5	19.1	2140	10	AF163665 Mus muscu
C 45	84.5	19.1	2398	10	AF282322 Mus muscu

ALIGNMENTS


```

Db 42 ATGAGGCTTCAGTCCCTTTCCAGCGCTGCTGTGATCTCTGCTTCTCTCTCCATCTTC 101
Qy 21 SerThrGluGlyLysArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 102 TCCACAGAGGAGAGAGCGCTCTGCCAAGCGCTGTGTACGAGCAGAGAACAGGCTCTGC 161
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys 60
Db 162 TGCCACCGAGTCCCTAGTCCCACTCAACAACTGAAAGGACATCATGTGAGGCTCTGT 221
Qy 61 LysProCysLysLeuGluProGluProArgLeuValProGlyAlaLeuProGln 80
Db 222 AACACGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 281
Qy 81 Val 81
Db 282 GTG 284

RESULT 4
AX342218
LOCUS AX342218 806 bp DNA linear PAT 12-JAN-2002
DEFINITION Sequence 4 from Patent WO0198497.
ACCESSION AX342218
VERSION AX342218.1 GI:18151763
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Polverino,A.J. and Luethy,R.
TITLE Secreted epithelial colon stromal-1 polypeptides, nucleic acids
encoding the same and uses thereof
JOURNAL Patent: WO 0198497-A 4 27-DEC-2001;
Angen, Inc. (US)
FEATURES
source 1. 806
/organism="Homo sapiens"
/db xref="taxon:9606"
29_274
/notes="unnamed protein product"
CDS
sig_peptide 207 a 257 c 179 g 163 t
BASE COUNT 207 a 257 c 179 g 163 t
ORIGIN
Alignment Scores:
Pred. No.: 2.75e-33 Length: 806
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-724-000A-5 (1-81) x AX342218 (1-806)
Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuLeuCysPheSerIlePhe 20
Db 29 ATGAGGCTTCAGTCCCTTTCCAGCGCTGCTGTGATCTCTCTCTCTCTCTCTCTCTTC 88
Qy 21 SerThrGluGlyLysArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 89 TCCACAGAGGAGAGAGCGCTCTGCCAAGCGCTGTGTACGAGCAGAGAACAGGCTCTGC 148
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys 60
Db 149 TGCCACCGAGTCCCTAGTCCCACTCAACAACTGAAAGGACATCATGTGAGGCTCTGT 208
Qy 61 LysProCysLysLeuGluProGluProArgLeuValProGlyAlaLeuProGln 80

```

```

Db 209 AAACCATGCAAGCTTGAGCCAGAGCCCGCTTTGGGTGGTGGCTGGGGCACTCCACAG 268
Qy 81 Val 81
Db 269 GTG 271

RESULT 5
AK025416
LOCUS AK025416 2063 bp mRNA linear PRI 29-SEP-2000
DEFINITION Homo sapiens cDNA: FLJ21763 fis, clone COLF6967.
ACCESSION AK025416
VERSION AK025416.1 GI:10437924
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens colon mucosa cDNA to mRNA, clone lib:ColF
clone:COLF6967.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H.,
Ota,T., Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2063)
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure Analysis, Human
Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639,
Japan (E-mail:cdna@ims.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- and 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source 1. 2063
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="COLF6967"
/tissue type="colon mucosa"
/clone_lib="ColF"
/notes="cloning vector pME18SFL3"
BASE COUNT 536 a 479 c 476 g 572 t
ORIGIN
Alignment Scores:
Pred. No.: 7.59e-33 Length: 2063
Score: 442.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-724-000A-5 (1-81) x AK025416 (1-2063)
Qy 1 MetArgLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuLeuCysPheSerIlePhe 20
Db 115 ATGAGGCTTCAGTCCCTTTCCAGCGCTGCTGTGATCTCTCTCTCTCTCTCTCTCTTC 174
Qy 21 SerThrGluGlyLysArgArgProAlaLysAlaTrpSerGlyArgArgThrArgLeuCys 40
Db 175 TCCACAGAGGAGAGAGCGCTCTGCCAAGCGCTGTGTACGAGCAGAGAACAGGCTCTGC 234
Qy 41 CysHisArgValProSerProAsnSerThrAsnLeuLysGlyHisValArgLeuCys 60
Db 235 TGCCACCGAGTCCCTAGTCCCACTCAACAACTGAAAGGACATCATGTGAGGCTCTGT 294

```



```

Db      279  GATA  282
RESULT 13
AC128725/c
LOCUS   AC128725
DEFINITION
Rattus norvegicus clone CH230-30014, *** SEQUENCING IN PROGRESS
***, 54 unordered pieces.
AC128725 137164 bp  DNA  linear  HTG 23-JUL-2002
AC128725.1 GI:21930188
HTG; HTGS PHASE1.
Rattus norvegicus.
Rattus norvegicus.
Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 137164)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homsai,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karissom,B., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newcomb,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviado,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojebokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatok,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs.R.
Direct Submission
Unpublished
2 (bases 1 to 137164)
Worley,K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSIT
Center clone name: CH230-30014
-----
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

----- Summary Statistics

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 81651 bases at least Q40
Consensus quality: 85938 bases at least Q30
Consensus quality: 89650 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 54 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1055: contig of 1055 bp in length
1056: gap of unknown length
1156: contig of 1098 bp in length
2253: gap of unknown length
2353: contig of 1482 bp in length
3835: gap of unknown length
5004: contig of 1069 bp in length
5104: gap of unknown length
7028: contig of 1923 bp in length
7127: gap of unknown length
8380: contig of 1253 bp in length
8480: gap of unknown length
9517: contig of 1037 bp in length
9617: gap of unknown length
11124: contig of 1507 bp in length
11224: gap of unknown length
12224: contig of 1000 bp in length
12324: gap of unknown length
13444: contig of 1120 bp in length
13544: gap of unknown length
14548: contig of 1004 bp in length
14648: gap of unknown length
16085: contig of 1417 bp in length
16165: gap of unknown length
17177: contig of 1012 bp in length
17277: gap of unknown length
18704: contig of 1427 bp in length
18804: gap of unknown length
19895: contig of 1091 bp in length
19995: gap of unknown length
21433: contig of 1438 bp in length
21533: gap of unknown length
22767: contig of 1234 bp in length
22867: gap of unknown length
24068: contig of 1201 bp in length
24168: gap of unknown length
25784: contig of 1616 bp in length
25884: gap of unknown length
27104: contig of 1220 bp in length
27204: gap of unknown length
28209: contig of 1005 bp in length
28309: gap of unknown length
30046: contig of 1737 bp in length
30146: gap of unknown length
31379: contig of 1233 bp in length
31479: gap of unknown length
33395: contig of 1916 bp in length
33495: gap of unknown length
35282: contig of 1787 bp in length
35382: gap of unknown length
38059: contig of 2677 bp in length
38159: gap of unknown length
40027: contig of 1868 bp in length
40127: gap of unknown length
42034: contig of 1907 bp in length


```

/note="assembly_fragment"
27230..42139
/note="assembly_fragment"
42240..59229
/note="assembly_fragment"
59330..78798
/note="assembly_fragment"
78899..144433
/note="assembly_fragment"
144534..225929
/note="assembly_fragment"
64591 t 900 others

BASE COUNT 67274 a 46621 c 46543 g 64591 t
ORIGIN

Alignment Scores:
Pred. No.: 108 Length: 225929
Score: 106.50 Matches: 24
Percent Similarity: 50.87% Conservative: 4
Best Local Similarity: 52.17% Mismatches: 15
Query Match: 24.10% Indels: 3
DB: 2 Gaps: 1

US-09-724-000A-5 (1-81) x AC109260 (1-225929)

Qy 1 MetArgLeuValLeuSerSerLeuLeuCySileLeuLeuCySPhSerIlePhe 20
Db 24196 ATGAGACTTCAGCCCTTCGGCTGCTCTGCATGCTGCTCTCTGTTTCGATTTC 24255
Qy 21 SerThrGlucIlyysArgArg-----ProAlaIysAlaTrpSerGlyAArgGThr 37
Db 24256 TCCTCAGAGAGTAACTCGCCCTGGCGCATTCGCTGCTGCTGGGAAGGTTTGAACAG 24315
Qy 38 ArgLeuCySPhisArg 43
Db 24316 AAAC TAGTGGGGCATAGG 24333

RESULT 15
CNS07EFU/c 207374 bp DNA linear PRI 06-JUL-2002
LOCUS Human chromosome 14 DNA sequence BAC R-982M15 of library RPCI-11
DEFINITION from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL583722
VERSION AL583722.5 GI:21712965
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 207374)
Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Catolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Eckenberg,R., Brula,T., deBerardinis,V., Cruaud,C.,
Cyapay,G., Saurin,W. and Weissensbach,J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 207374)
Genoscope.
Direct Submission
Submitted (05-JUL-2002) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE [E-mail : seqref@genoscope.cns.fr]
- Web : www.genoscope.cns.fr)
On Jul 8, 2002 this sequence version replaced gi:15526346.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
----- Finishing boundaries
FINISHED SEGMENT STARTS AT BASE 34250
FINISHED SEGMENT ENDS AT BASE 211616
-----

```

----- Summary Statistics
 Assembly program: Phrap; version 2.0

 Overall quality chart :

Range : bases

0 : bases

1 - 9 : 8

10 - 19 : 52

20 - 29 : 194

30 - 39 : 888

40 - 49 : 5766

50 - 59 : 10291

60 - 69 : 11885

70 - 79 : 24619

80 - 89 : 59705

90 - 99 : 93966

Percentage of bases with a quality value >= 40 : 99 %.

FEATURES

SOURCE

Location/Qualifiers

1. .207374

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="14"

/clone="R-982M15"

/clone_lib="RPCI-11"

48734..50176

/evidence=not_experimental

/insertion_seq="IS150"

123536..123658

/note="matching EMBL:AA431297

RHdb:RH98436

dbSTS:STS68194

Identified using the e-PCR software (G. Schuler)"

127693..127862

/note="matching EMBL:AA621795

RHdb:RH102965

dbSTS:STS70502

Identified using the e-PCR software (G. Schuler)"

155522..155646

/note="matching EMBL:G19662

RHdb:RH10471

dbSTS:STS13897

Identified using the e-PCR software (G. Schuler)"

155534..155683

/note="matching EMBL:H82735

RHdb:RH53711

dbSTS:STS8268

Identified using the e-PCR software (G. Schuler)"

177768..177923

/note="matching EMBL:AA056416

RHdb:RH47592

dbSTS:STS40651

Identified using the e-PCR software (G. Schuler)"

177906..178028

/note="matching EMBL:AA504155

RHdb:RH101916

dbSTS:STS69453

Identified using the e-PCR software (G. Schuler)"

177971..178228

/note="matching EMBL:M63167

RHdb:RH69160

dbSTS:STS49050

Identified using the e-PCR software (G. Schuler)"

46690 a 60047 c 58929 g 41708 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 2.37e+03

Score: 22.00

Percent Similarity: 50.00%

Best Local Similarity: 38.57%

Query Match: 20.81%

Length: 207374

Matches: 27

Conservative: 8

Mismatches: 29

Indels: 6

DB: 9 Gaps: 3

US-09-724-000A-5 (1-81) x CNS07EFU (1-207374)

QY 3 LeuLeuValLeuSerSerLeuLeuCysIleLeuLeuLeuCysPheSerIlePheSerThr 22
 |||||
 Db 130256 CTGTTGGTTTGGTTCTCTTTGCTGGCTGTAGCCCACTAAACTGACTTCTCATGGAGCACA 130197
 |||||

QY 23 GluGlyLysArgArgProAlaLysAlaIlePheSerGlyArgArgThrArgLeuCysHis 42
 |||||
 Db 130196 GAGGAAAAAGAAAAACCA-----GGCTCCAGGAGGGCAGCTCTCTCTGTCTAC 130149
 |||||

QY 43 ArgValProSerProAsnSerThrAsnLeuLysGlyHisHisValArgLeuCysLys--- 61
 |||||
 Db 130148 AGGTCCTCCAGTCCCACTCTCCAGGCAGGAGGGCGCATGGTCTCACTTGTCTGCAGATGA 130089
 |||||

QY 62 ProCysLysLeuGluProGluProArgLeu 71
 |||||

Db 130088 CCTTGTAGG---TGACCCCACTCAATCTC 130062

Search completed: May 17, 2003, 11:55:39

Job time : 2310 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 16, 2003, 15:27:44 ; Search time 39 Seconds
(without alignments)
276.751 Million cell updates/sec

Title: US-09-724-000A-5

Perfect score: 81

Sequence: 1 MRLVLVSLLCILLCSIF.....PCKLEPEPRLWVGPALPQV 81

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 229602

Minimum DB seq length: 23

Maximum DB seq length: 80

Post-processing: Listing first 45 summaries

Database : A_Geneseq101002.*
1: /SID82/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID82/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID82/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID82/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID82/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID82/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID82/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID82/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID82/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID82/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID82/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID82/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID82/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID82/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID82/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID82/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID82/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID82/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID82/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID82/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID82/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID82/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID82/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	57	70.4	57	AAE16482	Human Secreted epi
2	30	37.0	30	AAE16485	Human Secreted epi
3	28	34.6	28	AAE16486	Human Secreted epi
4	23	28.4	23	AAE16484	Human Secreted epi
5	9	11.1	54	AAE16480	Mouse Secreted epi
6	9	11.1	77	AAE16483	Rat Secreted epi
7	9	11.1	78	AAE16479	Mouse Secreted epi
8	7	8.6	72	AAE19500	Amino acid sequenc
9	7	8.6	78	AAU45851	Propionibacterium
10	6	7.4	25	AAB39415	Human secreted pro

11	6	7.4	28	AAU05966	Cone snail O-supe
12	6	7.4	31	AAW81050	Signal peptide-cyt
13	6	7.4	32	AAW92181	Human digestive sy
14	6	7.4	39	AAAP30235	Sequence of cortic
15	6	7.4	47	ABP41217	Human ovarian anti
16	6	7.4	48	ABP33769	Human secreted pro
17	6	7.4	52	AAW56482	Arabidopsis thalia
18	6	7.4	52	AAW58744	Arabidopsis thalia
19	6	7.4	52	AAW53929	Propionibacterium
20	6	7.4	52	AAU56623	Propionibacterium
21	6	7.4	53	ABP33245	Human ORF2218 prot
22	6	7.4	53	ABP34966	Human ORF3939 prot
23	6	7.4	54	AAW48477	Human breast tumou
24	6	7.4	54	AAW72480	Human bone marrow
25	6	7.4	54	AAW32728	Peptide #6765 enco
26	6	7.4	54	ABG42300	Human peptide enco
27	6	7.4	55	AAV91504	Human secreted pro
28	6	7.4	56	AAU55581	Propionibacterium
29	6	7.4	56	ABP07189	Human ORFX protein
30	6	7.4	58	AAW85937	Human immune/haema
31	6	7.4	59	ABP31870	Human ORF843 prote
32	6	7.4	65	AAU22012	Human cardiovascular
33	6	7.4	68	AAW09986	Arabidopsis thalia
34	6	7.4	69	ABW22742	Escherichia coli p
35	6	7.4	69	ABP09715	Human ORFX protein
36	6	7.4	70	AAW25064	Plant SDF encoded
37	6	7.4	72	ABW30289	Peptide #2940 enco
38	6	7.4	72	ABW35455	Peptide #2961 enco
39	6	7.4	72	ABW20895	Protein #2894 enco
40	6	7.4	72	AAW56278	Human brain expres
41	6	7.4	72	AAW68653	Human bone marrow
42	6	7.4	72	AAW16469	Peptide #2903 enco
43	6	7.4	72	AAW28962	Peptide #2999 enco
44	6	7.4	72	AAW04197	Peptide #2879 enco
45	6	7.4	72	ABG38233	Human peptide enco

ALIGNMENTS

RESULT 1

AAE16482

ID AAE16482 standard; peptide; 57 AA.

AC AAE16482;

DT 09-APR-2002 (first entry)

DE Human Secreted epithelial colon stromal-1 (Secs-1) protein fragment.

Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
human.

OS Homo sapiens.

XX WO200198497-A1.

XX 27-DEC-2001.

XX 28-NOV-2000; 2000WO-US32479.

XX 21-JUN-2000; 2000US-0599087.

XX 28-NOV-2000; 2000US-0724000.

XX (AMGE-) AMGEN INC.

XX Polverino AJ, Luethy R;

XX

DR WPI; 2002-122281/16.
 XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids.
 PT useful for diagnosing, treating and preventing hematopoietic disorder.
 PT osteoporosis, Paget's disease, cancer, diabetes -
 XX Claim 14; Page 122; 134pp; English.
 XX The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as hematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is human Secs-1
 CC protein fragment.
 XX SQ Sequence 57 AA;
 Query Match 70.4%; Score 57; DB 23; Length 57;
 Best Local Similarity 100.0%; Pred. No. 6.9e-50;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 25 KRRPAKAWSGRRTRLCCHRVSPNSTNLKGHHVRLCKPCKLEPEPRLWVPGALPOV 81
 |||||
 DB 1 KRRPAKAWSGRRTRLCCHRVSPNSTNLKGHHVRLCKPCKLEPEPRLWVPGALPOV 57
 RESULT 2
 AAE16485
 ID AAE16485 standard; peptide; 30 AA.
 AC AAE16485;
 XX 09-APR-2002 (first entry)
 DT Human Secreted epithelial colon stromal-1 (Secs-1) peptide #2.
 DE Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW hematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
 KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
 KW human.
 XX Homo sapiens.
 OS WO200198497-A1.
 XX 27-DEC-2001.
 XX 28-NOV-2000; 2000WO-US32479.
 XX 21-JUN-2000; 2000US-0599087.
 XX 28-NOV-2000; 2000US-0724000.
 XX (AMGE-) AMGEN INC.
 XX Polverino AJ, Luethy R;
 XX WPI; 2002-122281/16.
 XX N-PSDB; AAD27026.
 XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 PT useful for diagnosing, treating and preventing hematopoietic disorder.

PT osteoporosis, Paget's disease, cancer, diabetes -
 XX Disclosure; Fig 4D; 134pp; English.
 XX The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as hematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is human Secs-1
 CC peptide encoded by second exon.
 XX SQ Sequence 30 AA;
 Query Match 37.0%; Score 30; DB 23; Length 30;
 Best Local Similarity 100.0%; Pred. No. 6.1e-23;
 Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 24 GRRPAKAWSGRRTRLCCHRVSPNSTNLK 53
 |||||
 DB 1 GRRPAKAWSGRRTRLCCHRVSPNSTNLK 30
 RESULT 3
 AAE16486
 ID AAE16486 standard; peptide; 28 AA.
 AC AAE16486;
 XX 09-APR-2002 (first entry)
 DT Human Secreted epithelial colon stromal-1 (Secs-1) peptide #3.
 DE Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW hematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
 KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
 KW human.
 XX Homo sapiens.
 OS WO200198497-A1.
 XX 27-DEC-2001.
 XX 28-NOV-2000; 2000WO-US32479.
 XX 21-JUN-2000; 2000US-0599087.
 XX 28-NOV-2000; 2000US-0724000.
 XX (AMGE-) AMGEN INC.
 XX Polverino AJ, Luethy R;
 XX WPI; 2002-122281/16.
 XX N-PSDB; AAD27026.
 XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 PT useful for diagnosing, treating and preventing hematopoietic disorder.
 PT osteoporosis, Paget's disease, cancer, diabetes -
 XX Disclosure; Fig 4F; 134pp; English.

CC The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as haematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is human Secs-1
 CC peptide encoded by third exon.

XX Sequence 28 AA;

Query Match 34.6%; Score 28; DB 23; Length 28;
 Best Local Similarity 100.0%; Pred. No. 6e-21;
 Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 54 GHVRLCKPCLEPEPRLWVPGALPOV 81
 Db 1 GHVRLCKPCLEPEPRLWVPGALPOV 28
 |||||

RESULT 4

AAE16484
 ID AAE16484 standard; peptide; 23 AA.

XX AAE16484;

XX 09-APR-2002 (first entry)

DE Human Secreted epithelial colon stromal-1 (Secs-1) peptide #1.

XX Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
 KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
 KW human.

XX Homo sapiens.

XX WO200198497-A1.

XX 27-DEC-2001.

XX 28-NOV-2000; 2000WO-US32479.

XX 21-JUN-2000; 2000US-0599087.

XX 28-NOV-2000; 2000US-0724000.

XX (AMGE-) AMGEN INC.

XX Polverino AJ, Luethy R;

XX WPI; 2002-122281/16.

XX N-PSDB; AAD27026.

XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 PT useful for diagnosing, treating and preventing hematopoietic disorder,
 PT osteoporosis, Paget's disease, cancer, diabetes -

XX Disclosure; Fig 4A; 134pp; English.

XX The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound

CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as haematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is human Secs-1
 CC peptide encoded by first exon.

XX Sequence 23 AA;

Query Match 28.4%; Score 23; DB 23; Length 23;
 Best Local Similarity 100.0%; Pred. No. 5.4e-16;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 MRLVLSSLLCILLCSIFSTE 23
 Db 1 MRLVLSSLLCILLCSIFSTE 23
 |||||

RESULT 5

AAE16480
 ID AAE16480 standard; peptide; 54 AA.

XX AAE16480;

XX 09-APR-2002 (first entry)

DE Mouse Secreted epithelial colon stromal-1 (Secs-1) protein fragment.

XX Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
 KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
 KW mouse.

XX Mus musculus.

XX WO200198497-A1.

XX 27-DEC-2001.

XX 28-NOV-2000; 2000WO-US32479.

XX 21-JUN-2000; 2000US-0599087.

XX 28-NOV-2000; 2000US-0724000.

XX (AMGE-) AMGEN INC.

XX Polverino AJ, Luethy R;

XX WPI; 2002-122281/16.

XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 PT useful for diagnosing, treating and preventing hematopoietic disorder,
 PT osteoporosis, Paget's disease, cancer, diabetes -

XX Claim 14; Page 120; 134pp; English.

XX The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as haematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1

CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is mouse Secs-1
 CC protein fragment.

XX Sequence 54 AA;

Query Match 11.1%; Score 9; DB 23; Length 54;
 Best Local Similarity 100.0%; Pred. No. 0.15;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 WVPGALPQ 80
 |||||
 Db 45 WVPGALPQ 53

RESULT 6
 AAE16483
 ID AAE16483 standard; Protein; 77 AA.

XX AAE16483;

XX 09-APR-2002 (first entry)

DE Rat Secreted epithelial colon stromal-1 (Secs-1) protein.

XX Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
 KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
 KW rat.

XX Rattus norvegicus.

OS WO200198497-A1.

XX 27-DEC-2001.

XX 28-NOV-2000; 2000WO-US32479.

XX 21-JUN-2000; 2000US-0599087.

PR 28-NOV-2000; 2000US-0724000.

XX (AMGE-) AMGEN INC.

XX Polverino AJ, Luethy R;

XX WPI; 2002-122281/16.

XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 PT useful for diagnosing, treating and preventing hematopoietic disorder,
 PT osteoporosis, Paget's disease, cancer, diabetes -

PS Disclosure; Fig 3; 134pp; English.

XX The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as haematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or
 CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is rat Secs-1

CC protein.

XX Sequence 77 AA;

Query Match 11.1%; Score 9; DB 23; Length 77;
 Best Local Similarity 100.0%; Pred. No. 0.21;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 WVPGALPQ 80
 |||||
 Db 68 WVPGALPQ 76

RESULT 7
 AAE16479
 ID AAE16479 standard; Protein; 78 AA.

XX AAE16479;

XX 09-APR-2002 (first entry)

XX Mouse Secreted epithelial colon stromal-1 (Secs-1) protein.

XX Secreted epithelial colon stromal-1; Secs-1; gene therapy; osteopetrosis;
 KW haematopoietic disorder; osteoporosis; osteogenesis imperfecta; cachexia;
 KW Paget's disease; periodontal disease; hypercalcaemia; glomerulonephritis;
 KW diabetes; obesity; osteopathic; cytostatic; nephrotrophic; antidiabetic;
 KW anorectic; immunomodulator; antipsoriatic; vulnery; antiinfertility;
 KW gynaecological; antiulcer; antiinflammatory; cancer; cell therapy;
 KW mouse.

XX Mus musculus.

XX Key Location/Qualifiers

FT Peptide 1..24

FT Protein /label= Signal_peptide

FT Protein 25..78

FT Protein /label= Mature_mouse_Secs-1_protein

PN WO200198497-A1.

XX 27-DEC-2001.

XX 28-NOV-2000; 2000WO-US32479.

XX 21-JUN-2000; 2000US-0599087.

PR 28-NOV-2000; 2000US-0724000.

XX (AMGE-) AMGEN INC.

XX Polverino AJ, Luethy R;

XX WPI; 2002-122281/16.

DR N-PSDB; AAD27024.

XX Secreted epithelial colon stromal-1 polypeptides and nucleic acids,
 PT useful for diagnosing, treating and preventing hematopoietic disorder,
 PT osteoporosis, Paget's disease, cancer, diabetes -

XX Claim 13; Fig 3; 134pp; English.

XX The present invention relates to an isolated murine or human secreted
 CC epithelial colon stromal-1 (Secs-1) polypeptide, its allelic or splice
 CC variant, orthologue, fragment or mutant. Secs-1 gene is used in gene
 CC therapy and cell therapy. Secs-1 is useful for identifying a compound
 CC which binds to a Secs-1 polypeptide. Secs-1 is useful for treating,
 CC preventing or ameliorating a disease condition such as haematopoietic
 CC disorder, osteoporosis, osteopetrosis, osteogenesis imperfecta, Paget's
 CC disease, periodontal disease, hypercalcaemia, acute glomerulonephritis,
 CC chronic glomerulonephritis, cancer, diabetes, obesity or cachexia. Secs-1
 CC is also useful for diagnosing a pathological condition which involves
 CC determining the presence or amount of Secs-1 or polypeptide encoded by
 CC Secs-1 DNA in a sample; and diagnosing a pathological condition, or

CC susceptibility to pathological condition based on the presence or amount
 CC of expression of the polypeptide. The present sequence is mouse Secs-1
 CC protein.

XX Sequence 78 AA;

Query Match 11.1%; Score 9; DB 23; Length 78;
 Best Local Similarity 100.0%; Pred. No. 0.21; Indels 0; Gaps 0;
 Matches 9; Conservative 0; Mismatches 0;

Oy 72 WVPGLPQ 80

Db 59 WVPGLPQ 77

RESULT 8

AAU19500
 ID AAY19500 standard; Protein; 72 AA.

XX AC AAY19500;

XX 14-JUL-1999 (first entry)

XX Amino acid sequence of a human secreted protein.

XX Human secreted protein; cancer; tumour; neurodegenerative disorder;
 KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
 KW immune system disease; autoimmune disease; hepatic disease;
 KW renal disease; lymphoma; inflammation; allergy; ischemic shock;
 KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
 KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
 KW lung disease; thymus disease; digestive disorder; endocrine disorder;
 KW infection; AIDS.

XX Homo sapiens.

XX WO9922243-A1.

XX 06-MAY-1999.

XX 23-OCT-1998; 98WO-US22376.

XX 24-OCT-1997; 97US-0063387.

XX 24-OCT-1997; 97US-0062784.

XX 24-OCT-1997; 97US-0063088.

XX 24-OCT-1997; 97US-0063089.

XX 24-OCT-1997; 97US-0063090.

XX 24-OCT-1997; 97US-0063091.

XX 24-OCT-1997; 97US-0063092.

XX 24-OCT-1997; 97US-0063097.

XX 24-OCT-1997; 97US-0063098.

XX 24-OCT-1997; 97US-0063099.

XX 24-OCT-1997; 97US-0063100.

XX 24-OCT-1997; 97US-0063101.

XX 24-OCT-1997; 97US-0063109.

XX 24-OCT-1997; 97US-0063110.

XX 24-OCT-1997; 97US-0063111.

XX 24-OCT-1997; 97US-0063114.

XX 24-OCT-1997; 97US-0063386.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;
 PI Feng P, Florence C, Florence KA, Greene JM, Janat F;
 PI Kayw H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;

XX WPI; 1999-303069/25.

XX N-PSDB; AAX61380.

XX New isolated human genes and the secreted polypeptides they encode

XX Claim 11; Page 416; 546pp; English.

PS

XX

CC The specification describes cDNA sequences (AAX61322-X61470) encoding
 CC human secreted proteins (AAY19442-Y19590). The polynucleotides and their
 CC corresponding secreted polypeptides are useful for preventing, treating
 CC or ameliorating medical conditions, e.g. by protein or gene therapy.
 CC Pathological conditions can also be diagnosed by determining the amount
 CC of the polypeptides in a sample or by determining the presence of
 CC mutations in the polynucleotides. Specific uses are described for each
 CC of the polynucleotides, based on which tissues they are most highly
 CC expressed in, and include developing products for the diagnosis or
 CC treatment of cancer, tumours, neurodegenerative disorders, developmental
 CC abnormalities and fetal deficiencies, blood disorders, leukemias, and
 CC diseases of the immune system, autoimmune diseases, hepatic and renal
 CC disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
 CC and cognitive disorders, schizophrenia, prostate diseases, obesity,
 CC disorders involving osteoclasts such as osteoporosis, arthritis or
 CC malignancies, diseases of testes, lung or thymus, digestive/endocrine
 CC disorders, infections and AIDS. The polypeptides are also useful for
 CC identifying their binding partners.

XX Sequence 72 AA;

Query Match 8.6%; Score 7; DB 20; Length 72;
 Best Local Similarity 100.0%; Pred. No. 21; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0;

Oy 13 LLLCFSI 19

Db 22 LLLCFSI 28

RESULT 9

AAU45851

ID AAU45851 standard; Protein; 78 AA.

XX AC AAU45851;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #6747.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US12865.

XX 21-APR-2000; 2000US-199047P.

XX 02-JUN-2000; 2000US-208841P.

XX 07-JUL-2000; 2000US-216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59528.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris -

XX Example 1; SEQ ID No 7046; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic

XX

CC

CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertrophia and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA).
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 78 AA;

Query Match 8.6%; Score 7; DB 22; Length 78;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 30 KAWSGRR 36
 |||||
 DB 55 KAWSGRR 61

RESULT 10

AAB39415
 ID AAB39415 standard; Protein; 25 AA.

AC AAB39415;

DT 02-FEB-2001 (first entry)

DE Human secreted protein sequence encoded by gene 14 SEQ ID NO:74.

XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW nontropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
 KW food additive; preservative; chromosome identification; cancer;
 KW female reproductive system disorder; immune disorder; wound healing;
 KW cardiovascular disorder; neurological disease; infectious disease;
 KW infection.

XX Homo sapiens.

OS WO200058340-A2.

PN 05-OCT-2000.

PD 23-MAR-2000; 2000WO-US07724.

PF 26-MAR-1999; 99US-0126510.

PR 07-JAN-2000; 2000US-0174850.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Rosen CA, Ruben SM, Komatsoulis G;

PI WPI; 2000-594638/56.

DR N-PSDB; AAC74350.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases -

XX Claim 11; Page 362; 391pp; English.

XX The polynucleotide sequences given in AAC74337 to AAC74386 encode the
 CC human secreted proteins given in AAB39402 to AAB39451. AAB39452 to
 CC AAB39484 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Example of activities include:
 CC cytostatic; immunosuppressive; nontropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;
 CC vulnery; anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for preventing,
 CC treating or ameliorating a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
 CC can also be used as a food additive or preservative to increase or
 CC decrease storage capabilities. The polynucleotide are useful for
 CC chromosome identification. They are also useful as probes for diagnosing
 CC a disorder related to the female reproductive system, particularly breast
 CC and/or ovary cancer. They are also useful in the gene therapy of breast
 CC and ovarian cancer. Secreted protein nucleic acids, proteins,
 CC antibodies, agonists and antagonists are useful in the diagnosis,
 CC treatment and prevention of: (a) cancer; (b) immune disorders; (c)
 CC cardiovascular disorders; (d) wound healing; (e) neurological diseases;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections. AAC74328 to AAC74336 and AAB39401 represent
 CC sequences used in the exemplification of the present invention.

XX Sequence 25 AA;

Query Match 7.4%; Score 6; DB 21; Length 25;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLSSLL 10
 |||||
 DB 11 VLSSLL 16

RESULT 11

AAU05966
 ID AAU05966 standard; Peptide; 28 AA.

XX AAU05966;

DT 24-OCT-2001 (first entry)

DE Cone snail O-superfamily conotoxin, De6.2.

XX Cone snail; O-superfamily conotoxin; sodium channel;
 KW demyelinating disease; multiple sclerosis; Huntington's disease;
 KW neuropathy; carpal tunnel syndrome; cardiovascular disorder;
 KW congestive heart failure; cancer; immunosuppression; epilepsy;
 KW asthma; ischaemia; stroke; pain.

XX Conus delessertii.

OS WO200149312-A2.

PN 12-JUL-2001.

PD 28-DEC-2000; 2000WO-US35431.

PF 30-DEC-1999; 99US-0173754.

PR 26-JUN-2000; 2000US-0214263.

PR 20-JUL-2000; 2000US-0219440.

PR 27-OCT-2000; 2000US-0243412.

XX (UTAH) UNIV UTAH RES FOUND.

PA (COGN-) COGNETIX INC.

XX Olivera BM, Cartier GE, Watkins M, Hillyard DR, McIntosh JM;

PI Layer RT, Jones RM;

XX WPI; 2001-418352/44.

XX New O-superfamily polypeptides useful for treating voltage gated ion
PT channel disorders, including demyelinating diseases i.e. multiple
PT sclerosis -
XX
PS Claim 2; Page 69; 277pp; English.
XX
CC The sequence is a cone snail O-superfamily conotoxin peptide.
CC The peptides are useful for regulating the flow of sodium through
CC sodium channels in an individual and the treatment or prevention of
CC disorders associated with voltage gated ion channel disorders,
CC including demyelinating diseases i.e. multiple sclerosis, optic
CC neuromyelitis, disseminated encephalomyelitis, adrenoleukodystrophy,
CC acute transverse myelitis, progressive multifocal leukoencephalopathy,
CC sub acute sclerosing panencephalomyelitis (SSPE), metachromatic
CC leukodystrophy, Pelizaeus-Werzbacher disease, spinal cord injury,
CC botulinum toxin poisoning, Huntington's, compression, entrapment
CC neuropathies i.e. ulnar nerve palsy, and carpal tunnel syndrome.
CC cardiovascular disorders, i.e. cardiac arrhythmias and congestive heart
CC failure, reactive gliosis, hyperglycaemia, immunosuppression, cocaine
CC addiction, cancer, cognitive dysfunction, neurotransmitter disorders
CC (i.e. Eaton-Lambert syndrome) and reversal of curare and other
CC neuromuscular blocking drugs. The neurological disorder is a seizure,
CC preferably one associated with epilepsy. The neurological disorder is a
CC neurotoxic injury associated with hypoxia, anoxia or ischaemia. The
CC neurotoxic injury is associated with stroke, cerebrovascular accident,
CC brain or spinal cord trauma, myocardial infarct, physical trauma,
CC drownings, suffocation, perinatal asphyxia or hypoglycaemic events. The
CC disorder is pain i.e. migraine, acute pain, persistent pain,
CC neuropathic pain, nociceptive pain. The disorder is inflammation or a
CC cardiovascular disorder. A conotoxin peptide of is useful to
CC alleviate pain in a mammal in pain or about to be subjected to a pain
CC causing event, and to treat disorders associated with radical
CC depolarisation of excitable membranes by activating a K_{ATP} channel, the
CC disorders include cardiac, ocular and cerebral ischaemia and asthma.
XX
SQ Sequence 28 AA;

Query Match 7.4%; Score 6; DB:22; Length 28;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 11 CILLIC 16
|||
Db 22 CILLIC 27
|||

RESULT 12
AAW81050
ID AAW81050 standard; Peptide; 31 AA.
XX
AC AAW81050;
XX
DT 10-MAY-1999 (first entry)
XX
DE Signal peptide-cytotoxic T cell epitope fusion.
XX
KW MG50; melanoma gene-50; melanoma associated antigen; human;
KW cancer; lung cancer; rhabdomyosarcoma; diagnosis; therapy; vaccine;
KW cytotoxic T cell epitope; signal peptide.
XX
OS Chimeric - Synthetic.
OS Chimeric - Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Peptide /note= "signal peptide"
FT Peptide 22..31
FT Peptide /note= "cytotoxic T cell epitope"
XX
PN WO9855133-A1.
XX
PD 10-DEC-1998.

XX 04-JUN-1998; 98WO-US11533.
XX
XX 06-JUN-1997; 97US-0870941.
XX
PA (REGC) UNIV CALIFORNIA.
PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX
XX Deans RJ, Kan-Mitchell J, Minev BR, Mitchell MG;
XX WPI; 1999-080820/07.
XX
XX New MG50 melanoma associated antigen fragments - used to develop
PT products for the detection, treatment and prevention of
PT MG50-expressing cancers, e.g. melanoma, lung cancer or
PT rhabdomyosarcoma
XX
PS Claim 12; Page 58; 79pp; English.
XX
CC This peptide comprises a claimed cytotoxic T cell epitope composed
CC of a signal peptide (see AAW81047) and a T cell epitope (see AAW81032)
CC derived from the human MG50 melanoma associated antigen (see
CC AAW81030). The signal peptide can facilitate presentation of the
CC epitope as a complex with an MHC molecule at the surface of an
CC antigen presenting cell (APC). The invention provides additional
CC T cell epitopes (see AAW81031-45) from MG50, and APCs that can
CC express an MG50 T cell epitope complexed with an MHC molecule on
CC their surfaces, and which can be administered to a patient having
CC a cancer that contains MG50-expressing cells in order to stimulate
CC an active immune response. MG50 vaccines comprising an MG50
CC polypeptide, a T cell epitope optionally attached to a signal
CC peptide, an anti-idiotypic antibody or a nucleic acid molecule
CC encoding an MG50 polypeptide or T cell epitope can be administered
CC for preventative or therapeutic purposes. The products and methods
CC are used for the detection, treatment and prevention of e.g.
CC melanoma, lung cancer and rhabdomyosarcoma.
XX
SQ Sequence 31 AA;

Query Match 7.4%; Score 6; DB:20; Length 31;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LLLCFS 18
|||
Db 11 LLLCFS 16
|||

RESULT 13
AAW92181
ID AAW92181 standard; Protein; 32 AA.
XX
AC AAW92181;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human digestive system antigen SEQ ID NO: 1530.
XX
KW Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum.
XX
OS Homo sapiens.
XX
PN WO200155314-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01324.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0228927.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 06-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-502630/55.
DR N-PSDB; AAK87954.
XX
PT Polynucleotides encoding digestive system antigens, useful for
PT diagnosing, treating, preventing and/or prognosing disorders of the
PT digestive system, particularly cancer and cancer metastases -
XX
PS Claim 11; SEQ ID NO 1530; 986pp; English.

XX The present invention provides the protein and coding sequences of a
 CC number of human digestive system antigens. These can be used in the
 CC diagnosis, treatment and prevention of digestive system disorders,
 CC including cancer, Meckel's diverticulum, bacterial or parasitic
 CC infections, appendicitis, Hirschsprung's disease, chronic colitis or
 CC ulcerative colitis. The present sequence is a digestive system antigen of
 CC the invention.

XX SQ Sequence 32 AA;

Query Match 7.4%; Score 6; DB 22; Length 32;

Best Local Similarity 100.0%; Pred. No. 99;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 13 LLLCFS 18

Db 13 LLLCFS 18

RESULT 14

AAP30235 ID AAP30235 standard; peptide; 39 AA.

AC AAP30235;

XX 25-APR-1992 (first entry)

XX Sequence of corticotropin, adrenocorticotrophic hormone (ACTH)
 DE analogue Tyr23-Phe2-Nle4-ACTH.

XX Hormone; pituitary; growth; adrenal gland.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 1 /label= H-Ser

FT Misc-difference 3 /label= Nle

FT Misc-difference 19 /label= Nle

FT Modified-site 39 /label= Nal

FT /label= Phe-OH

XX US4415546-A.

XX 15-NOV-1983.

XX 16-MAY-1983; 83US-0302746.

XX 12-MAY-1981; 81US-0262974.

XX (RAMA/) RAMACHANDRAN J.

XX Ramachandran J, Buckley DL, Yamashiro DH, Hagman JR;

XX WPI; 1983-832889/48.

XX ACTH analogues and their radioiodinated derivs. - useful as

FT corticosterone stimulators and in radioimmunoassays

XX Claim 1; Col 10; 6pp; English.

XX The ACTH analogues have biological activity fully comparable to that

CC of native ACTH and so are useful for stimulation of corticosterones.

CC The radioiodinated derivs. are useful in radioimmunoassays, and in

CC investigations of body processes, and unlike radioiodinated ACTH,

CC they do not show activity loss.

XX SQ Sequence 39 AA;

Query Match

7.4%; Score 6; DB 4; Length 39;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 25 KRRPAK 30

Db 16 KRRPAK 21

RESULT 15

ABP41217 ID ABP41217 standard; Protein; 47 AA.

AC ABP41217;

XX 23-AUG-2002 (first entry)

XX Human ovarian antigen HOCPO31, SEQ ID NO:2349.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;

KW ovarian cancer; breast cancer; reproductive system disorder;

KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;

KW inflammatory condition; immune disorder; blood disorder;

KW cardiovascular disorder; respiratory disorder; neurological disorder;

KW gastrointestinal disorder; urinary system disorder; drug screening;

KW gene therapy; chromosome mapping; forensic analysis;

KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;

KW antiinflammatory; gynaecological; reproductive.

XX Homo sapiens.

OS WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US18569.

XX 07-JUN-2000; 2000US-209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX N-PSDB; ABQ54294.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,

XX useful in the prevention, treatment and diagnosis of cancer (e.g.

XX ovarian cancer), immune disorders, cardiovascular disorders and

XX neurological diseases -

XX Claim 11; SEQ ID NO 2349; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-

CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also

CC encompasses polypeptides 90% identical and polynucleotides 95% identical

CC to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen

CC polynucleotides, antibodies against human ovarian antigens, and the use

CC of ovarian antigen polynucleotides and polypeptides in diagnosing,

CC treating, prognosing or preventing various ovary and/or breast-related

CC disorders. Such conditions include ovarian cancer and breast cancer, and

CC metastatic tumours of ovarian or breast origin, reproductive system

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,

CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine

CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic

CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and

CC vaginitis), immune disorders (e.g., congenital and acquired

CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),

CC blood-related disorders (e.g., anaemia), cardiovascular disorders,

CC respiratory disorders, neurological disorders, gastrointestinal disorders

CC and urinary system disorders. Ovarian antigen polypeptides and

CC polynucleotides may also be used in screening for compounds which

CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 47 AA;

Query Match 7.4%; Score 6; DB 23; Length 47;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LSSLIC 11
 |||||
 Db 6 LSSLIC 11

Search completed: May 16, 2003, 15:34:06
 Job time : 41 secs

73 x 2